Zorginstituut Nederland

PERSONAL HEALTH TRAIN

The Personal Health Train in health care

Stories from the work field

The **Personal Health Train** *(PHT) is designed to enable health care innovators and researchers to work with health data from various sources. It can give controlled access to data, while ensuring privacy protection and optimal engagement of individual patients and citizens.*

Table of contents

Introduction

This collection of stories from the work field shows how Dutch organisations use the Personal Health Train (PHT). FAIR data and the PHT are concepts that are being applied more and more often in order to re-use health-related data. Zorginstituut Nederland and Health-RI support this development and therefore collected these examples of divers practical Personal Health Train applications at the beginning of 2020. In order to inspire and to show the ongoing developments and possibilities. For even more examples of PHT-applications, visit the website of the PHT-network:www.personalhealthtrain.nl.

The Personal Health Train
Personal Health Train (PHT) is a collective term for technologies that enable the use and combination of data from different sources. The most important principle of the PHT is that the analysis is brought to the data, allowing the data to remain at the source. Keeping data at the source allows for privacy-by-design. The PHT enables the use and combination of the increasing amount of data from various data sources. The PHT facilitates source holders to set conditions and maintain control over the (purpose of the) use of the data. The expected benefits of the PHT principle have gained interest from various parties.

The purpose of this collection of examples
Many organisations are interested in the PHT or are already applying it. Thanks to the PHT network, they can find each other and brainstorm together about developments related to the Personal Health Train. During the Health-RI conference of 30 January 2020, the PHT network was officially launched and the associated website was officially presented. Organisations that are interested in the Personal Health Train can register for the network on the website and find information.

This collection of examples also contributes to the provision of information. These stories from the work field describes eleven organisations' experiences with the PHT. These show the current state of play and illustrate the broad range of application possibilities for the PHT. By this we intended to encourage collaboration between parties and also to learn from the tools and solutions that have been used.

The Personal Health Train-metaphor

In presentations and conversations about the Personal Health Train (PHT), people often use a train metaphor. The following explanation illustrates the different components of the PHT using this metaphor.

FAIR Data
The datasets must be described according to the FAIR principle (Findable, Accessible, Interoperable, Reusable) by means of a common ontology. An ontology is a computer-interpreted description (model) of reality, also known as a knowledge representation.

FAIR Data Point
The location where the owners of the data records their data is called a FAIR Data Point. In general, organisations make a selection of their (operational) data available at their data point and provide the data according to the FAIR Data principles of an ontology. Placing data in a FAIR Data Point does not mean that this data is automatically available to everyone.

FAIR Data Station
The FAIR Data Station facilitates that the train can access the data. Depending on the application, the service must be more or less secured; and depending on the complexity of the analysis, the service must offer more or fewer possibilities. This can range from an open SPARQL endpoint to a virtual machine.

Train
The train is the logic of the analysis. The form of the train varies, depending on the type of service. The train can take the form of a query, an interface call or a fully programmed and self-learning algorithm.

Rails
The rails are the agreements, guarantees and interfaces of the PHT.

Dashboard
The analysis is sent from a dashboard. This location has a monitor, which can track the progress of the train and show the final results.

1 Health–RI

Highlights

- Health-RI recognizes the use and importance of the Personal Health Train since 2014: It enables researchers, healthcare professionals and citizens to (re-)use health data to answer health care questions.
- Health-RI wants to facilitate and stimulate the PHT-development and its scale up, within the PHT community among others.
  An infrastructure for health data may support the scale-up. But many steps still need to be taken to link the technology to daily care practice.

Leone Flikweert

CEO Health-RI Foundation

Health-RI

Leone Flikweert is the CEO of the Health-RI Foundation. This is a public-private partnership of organisations involved in health research and health care. Health-RI is working on the development of an integrated health data (research) infrastructure that promotes the (re)use of knowledge, tools, facilities and health data, and enables a learning healthcare system. This can accelerate sustainable and affordable developments in personalised medicine and health. These ambitions form the basis for Health-RI's involvement in developments related to the Personal Health Train (PHT). Health-RI wants to facilitate, stimulate and scale up the development of the PHT. In this role, Health-RI has an overarching programme management function within the PHT programme.

The role of Health-RI

Health-RI wants to support the development of the PHT because it believes in the 'data at the source' approach that brings analysis to the data source. In the future, it sees data being stored in a more decentralised way in combination with federated data analyses. In addition, Health-RI sees opportunities to accelerate the development of the PHT. The development of the PHT started in the academic world, but has since generated a lot of external interest. It is not easy to branch out from the academic environment to routine use in daily care processes. According to Health-RI, the existence of a health data infrastructure can support the upscaling. Getting an adequate infrastructure still requires many steps in the link between technology and the daily care practice. Health-RI wants to further promote this development as a connecting factor between the various initiatives related to the PHT. The cooperating parties should work together on the PHT with the aim of improving and accelerating its development. In doing so, Health-RI also creates a collective voice for things that need to be changed at a government or legislative/regulatory level. To that end, Health-RI can conduct meetings with different parties on behalf of the collective. To stimulate the development of the PHT, Health-RI

places already developed tools on its website; the organisation recently also started publishing on the Personal Health Train website. The role of Health-RI in this is to provide an overview of what is and what isn't possible. In addition, Health-RI also has a facilitating role. With this facilitating role, Health-RI aims to support the development of the PHT, among other things by helping to develop a set of agreements.

The aim of the Personal Health Train

For Health-RI, the aim of the PHT is that every researcher, healthcare professional or citizen in the Netherlands has access to (their own) health data and has the opportunity to ask questions about this or to make analyses. This is a pipe dream for now. For the time being, Health-RI wants to use the PHT to bring about a situation in which all health data is accessible to researchers and healthcare providers, and by this means bring data closer together.

Personal Health Train developments

Health-RI recognizes that developments are needed in different areas. First of all, at a technological level. PHT developments are happening fast, but agreements need to be made to speed up the process. This includes standardisation, for example. Flikweert expects that market parties will play a major role in this, given the international interest in the subject. Big Data and AI are not just popular subjects in the Netherlands, but also abroad. For example, organisations are working hard on the development of health data infrastructures. Leone Flikweert: "In France, a major initiative has been developed: the health data hub. Another important initiative has been started in Germany, called the Medizin Informatik Initiative. These programmes have been initiated by the governments, to facilitate which agreements and provisions are needed to make health data accessible." Leone Flikweert expects obstacles to the development in the form of ethical, legal and organisational issues. These issues will vary from country to country, also within Europe, and this creates additional challenges. Many organisations in the Netherlands are struggling with questions such as: "Who has access to 'your data'?" And "When should permission be requested?" Health-RI has noticed the need for rules and clarity to scale the reuse of data for health purposes and research. Flikweert: "We notice in particular that cooperation and pooling of forces are needed." One of the organisational challenges is the question of trust. To make data accessible and adhere to the underlying agreements and conditions, trust within the network is essential. Currently, the various PHT networks are taking care of this themselves, but we need to work on scalable solutions.

> "Building up trust in person is not a scalable option.
> Standards and general agreements are needed to make sure everyone knows where they stand."

Development challenges

One of the obstacles to the development of the PHT is the funding of networks. The PHT method looks promising, but the business case is not always as clear. Funding for pilot schemes could be useful. Leone Flikweert: "We would like to set up some pilots where a PHT network is set up separately and where elements such as an agreement system can be tried out." Those pilots would also provide a nice opportunity to test the technology and link it to daily practice. Health-RI would like to close the gap between the various projects and the scalability.

In addition, ethical, legal and organisational aspects remain a challenge. At the moment, there are still many questions and it is not clear what is or is not possible. According to Flikweert, it is important that this be clarified, for example by mapping out the specific issues that can be answered. Answering these questions requires multidisciplinary collaboration. Health-RI would like to be part of the puzzle and work with the different parties to solve these issues.

Collaboration with other parties

In its coordinating role, Health-RI is working with many organisations in the field. Among other parties, Health-RI is working with the seven University Medical Centres (UMCs) to stimulate the development of the PHT. Health-RI also collaborates with those at the forefront of PHT development: MAASTRO, TNO, IKNL, Vektis, the Zorginstituut, CZ and ZN. Since 2019, a number of these parties have participated in the PHT coalition. Together with the coalition, an inventory is made of whether there is a need for an agreement system and how it can be implemented.

Coming year

Over the next two years, Health-RI and the coalition partners want to develop a PHT set of agreements. This also includes funding, which would help bring about the application of the Personal Health Train for the field. It is important to clarify what is and is not possible with the PHT at the moment. Next year, Leone Flikweert intends to develop a joint vision for the roadmap of the PHT development in the Netherlands.

Links

- Health-RI website
- PHT website
- Health Data Hub
- Medizin Informatik Initiative

2 Erasmus Medical Centre

Highlights

- The Erasmus MC wants to apply the PHT, including distributed learning, so that we can optimally learn from multicentric data.
- With other MCs, the Erasmus MC is working on Distributed and Federated Learning developments that allow software to run on local servers, enabling the same analysis to be run at different centres.
- During the first projects, the image data was analysed and aggregated in the same way at multiple MCs.
- The next step is to apply distributed deep learning, which will be used to train a neural network.
- To successfully use these techniques, we need to work toward an infrastructure that sets a standard for image data storage and access to it.

Wiro Niessen

Professor Biomedical Image Analysis, Erasmus MC

TU Delft

Wiro Niessen is a professor of Biomedical Image Analysis at Erasmus MC and as the CTO he is part of the Health-RI board. In recent years, the field of medical imaging has shown increasing interest in distributed learning on multicentric data for the development of AI techniques on large amounts of data. Image data is a valuable source for gaining insights into disease processes, forecasting, diagnostics, and treatment support. Medical imaging is therefore one of the largest data sources in daily medical practice and clinical research. Distributed learning is learning from data that is located at different places. This is interesting for analyses of large amounts of data such as image data without centralising the data. The PHT-principle enables us to do this.

Distributed learning and the Personal Health Train

The vision of Wiro Niessen's research group is to learn image data and thus improve health care. Image data is stored in different places and is difficult to share or centralise. That is why Erasmus MC has started to explore distributed learning, federated learning and multicentric learning, with the aim of learning more easily from different image data sources. Distributed and Federated Learning allow software to run on local servers, enabling the same analysis to be run at different centres. These techniques allow analysis to be performed on multicentric data, without the need to bring this data together. The concept of the PHT allows the software and analyses to be brought to different centres. Until now, this method of analysis has been developed by the Erasmus MC with other UMCs, because it mainly uses data from clinical or population studies. Eventually, the

Erasmus MC also wants to include clinical care data and image data from peripheral hospitals. For example, in several centres, MRI brain scans have been analysed in exactly the same way in the initial projects and the results have been combined. The next step is to apply distributed deep learning where a neural network is trained to perform a particular task on data stored in different centres. This way of learning can provide many new insights for diagnostic and prognostic purposes, among others. Wiro Niessen's research group will apply this to improve diagnosis in oncology and neurology.

Current state of affairs and vision

Erasmus MC has conducted a number of pilot studies and Proof of Concepts. One of these pilot studies was conducted in collaboration with Radboud University Medical Centre, University Medical Centre Utrecht, and LUMC. The same image analysis algorithms were performed in these hospitals. By running the same analyses in different centres, a multicentric data analysis could be carried out in exactly the same analysis environment, without the data being brought together. These projects are successful because all data is specifically prepared for them. That is why Niessen wants to work toward an infrastructure that sets a standard for image data storage and access to it. Such an infrastructure makes it possible to set out different types of questions, which can be answered because the different centres have standardised outcome measures. Erasmus MC's vision is to apply the PHT, including distributed learning, so that we can optimally learn from multicentric data. Wiro Niessen: "The reality is that data is divided between different places and thus is difficult to share and difficult to centralise. As soon as it becomes possible to analyse image data locally and combine the results, we can use the advanced processing power of computers. For example, this can be in the form of a neural network, while the data remains in different places. This can have a huge impact on the quality of diagnostics and forecasting in daily care and the opportunities for research."

The use of the Personal Health Train

To use image data for better diagnostics and forecasting, it is essential that centres make the image data in their archives accessible and thus available for retrieval. This can be achieved by setting up a data storage facility containing anonymised image data. This makes it possible to analyse data in different centres according to the concept of the Personal Health Train. To achieve this, discussions are also taking place with the Dutch Association for Radiology. The ultimate goal is that all hospitals anonymise and store their clinical image data so that this data can be used to answer various questions. But also that the analysis software can be applied and validated in the various places.

Development challenges of the Personal Health Train

The biggest challenge is the application of the FAIR principle to make the data available. It takes a lot of time to make data FAIR. It is important that different FAIR-described data sets are properly interconnected and mean the same thing. The semantics have to be the same. If the semantics are different and the same terms mean different things, it is not possible to automatically learn from previous outcomes in health care for future patients. Making data FAIR takes time and is a challenge, but it is important and necessary. The quality and efficiency gains are worth this investment. One point of attention is the generalisability of the algorithms and analyses. Analyses learn a pattern in data that trains them. It is desirable to have as much data as possible so that an algorithm learns based on heterogeneity. Just the generalisability is already a challenge. Data must be stored in a certain way before it can be used for learning algorithms. The PHT is already a good step in the right direction, as it allows algorithms to be tested in a multicentric setting.

Wiro Niessen is very enthusiastic about these developments and believes that distributed learning and artificial intelligence (AI) will play an increasingly important role. On the other hand, he realises that starting this movement requires a change in behaviour and culture from everyone in health care. To achieve this, it is important that everyone starts working in a data-driven manner in the interests of the patient and for better medicine. It is also important that we, as researchers, continue to demonstrate the added value of this data-supported approach. The PHT platform www.personalhealthtrain.nl was established for this purpose. Health-RI believes it is important to support the PHT initiative.

"We would like to use the information and images to develop better diagnostics and forecasting.
We learn from examples, clinical practice and research.
The idea is to learn from multicentric data to develop new diagnostics and forecasting, and deploy these in health care."

Plans for next year

The focus is on distributed deep learning because the Erasmus MC is keen to gain new insights into the relationship between various sciences and image data. Erasmus MC would also like to develop better diagnostic and forecasting algorithms by training the algorithms using data from multiple centres. It is also important to investigate how these techniques can be used in clinical practice. A first step is to develop a framework enabling the system to learn from multicentric data. Algorithms can then be validated and their added value determined. A second step is to identify techniques that are sufficiently robust and accurate to provide added value in clinical practice. This way of working - applying the PHT and distributed learning - can slowly lead to improvement in clinical practice through the use of data.

Links

ErasmusMC
Health-RI website

3 The Netherlands Comprehensive Organisation (IKNL)

Highlights

- Since 2017, IKNL has been applying three basic principles for the use of the Personal Health Train: the separation of trains, rails and stations during development, the use of open source, and sharing developments with the PHT network.
- IKNL has specialised in the development of the rails, an infrastructure called Vantage6, in collaboration with MAASTRO.
- The IKNL offers these products through open source.
- For this purpose, they have created a blog, are maintaining a source code Github page and have set up an online community where developers and data station administrators can ask questions about the technology (such as installing software and further train development).

Gijs Geleijnse
Senior Clinical Data Scientist & Programme Manager for Personal Health Train
Integraal Kankercentrum Nederland (IKNL)
(Comprehensive Cancer Organisation, IKNL)

Gijs Geleijnse works at the Comprehensive Cancer Centres (IKNL). The IKNL manages the Dutch Cancer Registry (NKR), which has been monitoring the care for cancer patients at a national level for over thirty years. IKNL has gathered a lot of scientifically based knowledge and insights that are valuable for the practice. These insights contribute to cancer prevention and help improve patient care and quality of life[1]. More and more questions call for the combination of NKR data with other sources, such as socio-economic data for patients. The Personal Health Train (PHT) enables the analysis of data from different sources. The IKNL has been studying this in depth since June 2017.

Using the Personal Health Train to learn from data

Gijs Geleijnse: "The Personal Health Train is a broad concept for different data analysis techniques, without the need to transfer data to third parties. One of these techniques is called federated learning". The PHT enables data analysis without third parties gaining access to the data or the data being traceable. The IKNL has invested in the PHT for three reasons. First, to enable other organisations or researchers to perform analyses on NKR datasets without the data leaving the IKNL. The second reason is that many research projects are interested in combining different data. Geleijnse: "Physicians and researchers are interested in knowing what the prognosis is for a cancer patient. To answer this question, data from the NKR can be enriched with data on treatments for other conditions. The PHT allows this data to be analysed and the question to be answered." The third reason is the collaboration with international organisations. The IKNL is part of various

[1] https://www.iknl.nl/over-iknl

international collaborations for cancer registration worldwide to identify cancer patterns. To recognise these patterns, all kinds of data are exchanged. The PHT provides a great solution to continue identifying these patterns, with maximum respect for patient privacy and without the need for centralizing data.

Rails, trains and stations

One of the main principles for IKNL is to separate the set-up of the rails, trains and stations and to develop them separately. This approach provides the opportunity to specialise in one of these techniques and to work together in connecting them. The IKNL specialises in the development of 'the rails'. In collaboration with MAASTRO, they have developed an open source infrastructure for the PHT. Gijs Geleijnse: "As far as we are concerned, the development of separate software components (i.e. the rails, the trains and the data stations) is a basic principle for separate - though interoperable - development. In this way, the code becomes more straightforward, the systems more transparent and therefore easier to understand and handle for people who want to use and reuse it. We call this 'separation of concerns'. According to Geleijnse, this approach is also strategically sound: "When these components are separated, problems can be solved separately, making it easier to understand. If something goes wrong somewhere, it's easier to locate where it went wrong." Because IKNL considers it important to develop open source products for the Personal Health Train, these products can be offered in a community. Within this community, it is then possible to contribute and build on the work of others. In turn, this contributes to the growth and success of the products. Gijs Geleijnse is convinced that the creation of a community that receives national and international contributions will make products more widely available and used. According to Geleijnse, it becomes possible to share everything with each other - except for the data itself. "If as many products as possible are shared and made public with the community, much can be achieved. By bringing together different backgrounds and perspectives, we can learn from each other and develop the best possible solutions. That is why we have created a blog on a website[2] with our partners, and maintain a source code Github page[3]. There is also an online community[4] where developers and administrators of data stations can ask questions about the technology (such as software installation and further development of trains)".

In addition, it is important for the IKNL to comply with the FAIR principles. They have the ambition to make their data station completely FAIR, but they also realise that they need to work hard to achieve this. This year, the IKNL has hired two data scientists with the aim of establishing the data station of the Dutch Cancer Registration according to the FAIR principles.

[2] Vantage6.ai

[3] https://github.com/iknl

[4] Accessible via vantage6.ai

Sharing infrastructure

IKNL has implemented the PHT on the basis of available software from MAASTRO: Vantage6 (six: Secure Insight eXchange). Gijs Geleijnse: "It is used in all research projects with the PHT and it is being continuously refined. The rails are therefore becoming more reliable. We have ensured that communication is done via encryption. This means that if a train is sent between two parties, the other parties cannot see what is in the package. For example, we have a working infrastructure for the rails that can be used in research projects. It is used in data analysis projects with our data; MAASTRO and Medical Data Works also use this in their collaborations. This infrastructure is documented; if others want to use it, the documentation and portal are available for setting up their own infrastructure. We now want to focus on creating more use cases."

"In the Netherlands, we have a huge amount of healthcare data that has enormous potential. If we together manage the Personal Health Train network well, I believe we can derive much more information from that data."

IKNL is also involved in the development of algorithms, the trains. These algorithms are being developed in collaboration with partners. IKNL has several partners who are working on solutions for both vertically and horizontally partitioned data [5] (including the Eindhoven University of Technology, the Netherlands Organisation for Applied Scientific Research (TNO) and Newcastle University). These trains are applied to epidemiological cancer research. One of these projects is an analysis based on data on the survival of cancer patients, in collaboration with the Taiwanese cancer registry. The technology has already been tested and a data request has been made, both in Taiwan and in the Netherlands. To this end, a PHT station has been installed in Taiwan to perform the analysis. At the Taiwanese cancer registry, it had been customary to grant access to the data through a system in which the data for cancer registrations was stored on a stand-alone computer that had no access to the Internet. This computer presented a shielded and therefore safe environment that only allowed access to people with special permission to perform an analysis. Following the collaboration with the IKNL and the confidence in the PHT approach, part of this data has been transferred to a computer with Internet access. Geleijnse: "Internet access is required if you want to use the PHT." A manuscript containing the results of this study has been sent to a clinical journal. Horizontal means that the patients are different, but the characteristics are the same.

Conducting more secure and frequent research

The IKNL hopes that the PHT can be used for many different types of cancer research and insights. With the help of the PHT, we will be able to answer major issues in the future. PHT will enable us to extract as much information as possible from data, with optimal protection for patient privacy. Gijs Geleijnse: "We therefore hope that the creation of a network of FAIR data stations will make the performance of research more accessible. This network makes cancer research more approachable because researchers can safely and easily ask the system a research question. "In the Netherlands, we have a huge amount of healthcare data that has enormous potential. I think that if we properly manage the PHT network, we can derive much more information from that data and can do much more to understand and combat cancer."

Points of special interest when developing the Personal Health Train

One of the things that researchers using the PHT face is the absence of a unified language. Gijs notes that it is extremely important to reach good agreements about how algorithms, the trains, are developed. For example, by developing a common language. The patient's gender is now recorded in many different ways. It would help to avoid errors to agree that this is always coded in one way. These agreements make it easier to conduct research.

[5] Vertical means that the patients are the same, but have different characteristics.

For the PHT, it is important to invest in these techniques. Another point that the IKNL has noticed is the need for more 'trains'. The IKNL has mainly implemented complex analyses, but retrospectively decided that simpler analyses, such as descriptive statistics, are also very useful information.

According to Gijs Geleijnse, the PHT still has a long way to go in terms of user-friendliness. "The PHT should be user-friendly, to increase usage. That is why we want to work toward the goal that the PHT can be used without the help of data analysts and experts in PHT technology." Because, according to Geleijnse, there is another challenge: "Implementing the PHT in organisations. Implementing the PHT also means that organisations need to align their processes with this technology. That means we need to move from data requests to Personal Health Train requests. That means not asking for the data that is needed, but delivering the analysis, which is sent to the data, and the answer is then returned."

Focus for the year 2020

The IKNL has two great ambitions for 2020. Firstly, they want to continue collaborating with TNO. IKNL is currently collaborating with TNO on a Multi-Party Computation (see also the Carla Rombouts interview). They are working on techniques from cryptography to analyse vertically partitioned data (so multiple organisations have different data from the same group of patients). The collaboration with TNO focuses on the technology.
A next step is to use these trains in a case that supports epidemiological research. IKNL's second ambition is to expand its impact in Europe. Little is known about rare cancers, because patient numbers are too small to perform statistical research. That is why it is necessary to step up our efforts at a European level. The IKNL would like to know what the impact of these tumours is, but also what good treatment patterns there are. In the next year, IKNL wants to use their network of European cancer registrations to set up studies that enable them to take a good look at cancer and the impact of cancer in Europe.

Collaboration with other parties

The IKNL works with the European Commission and its Joint Research Centre (JRC). The JRC shows, on the basis of cancer registrations throughout Europe, how many new diagnoses are made each year and what the differences are between European countries. For example, whether there is more prostate cancer in Slovakia than in Portugal. The IKNL is now working on a project to make all European cancer data available for the Personal Health Train in the long term. The second collaboration is with PALGA (the Dutch Pathological-Anatomical National Automated Archive), the national pathology platform that is carrying out a project on vertical petitioned data. This is a first pilot to test the technology. The goal is to implement a FAIR data station at PALGA, so that the PHT can be used.

Links

- The IKNL
- Github IKNL
- Vantage website and blog

4 Maastro Clinic

Highlights

- Has been working on the Personal Health Train since 2005
- Currently collaborates with 25 international centres, all connected by the Personal Health Train
- Has developed two infrastructures, a commercial one with Varian Medical Systems and an open-source, non-commercial infrastructure with IKNL
- MAASTRO uses the Personal Health Train for personalised medicine and shared decision making
- MAASTRO is developing tooling to facilitate the access data from electronic patient records.

André Dekker

Medical Physicist, Professor of Clinical Data Science

Maastro Clinic, Maastricht UMC,
Maastricht University
Co-lead of the international
PHT IN

André Dekker works at MAASTRO, a nationally and internationally renowned radiotherapy centre that emphasizes the connection between patient care, education and effective scientific research. Its vision: 'Ensuring oncology treatment with the greatest chance of recovery and/or as few physical and mental side effects as possible'[6]! This perspective is part of the basis for the interest in the Personal Health Train (PHT), which was generated 15 years ago. MAASTRO now collaborates 25 national and international centres; trains 'travel' daily between these centres, using distributed learning.

15 years of Personal Health Train

According to MAASTRO, PHT is a great application because the data stays at the source, but it can still be analysed. An important principle in this respect is that the source has total control and can also decide which question they want to answer. The stations must accept a question. Fifteen years ago, MAASTRO started to explore distributed learning; in 2015, the idea of the PHT was born. The collaboration with 25 centres was made possible because these centres were all applying the FAIR principles. MAASTRO has also developed two infrastructures for the PHT: a commercial one with Varian Medical Systems and an open-source, non-commercial infrastructure with IKNL. André Dekker: "We collaborate in particular with cancer hospitals worldwide. We use the PHT to learn which treatment, prescribed to which patient, will have which outcome. For this, we need a lot of data, from different parties. They input that data into the station and then the PHT can come by and perform the analyses."

MAASTRO uses the PHT for personalised medicine and shared decision making, also called

[6] Maastro Vision

joint decision making. This is possible by teaching algorithms what the outcomes of therapies are. Preventive models are created so that they can be used for joint decision making. In practice, this is used to inform patients about the success rate for treatments and the expected outcomes. Experience shows that there is still a lot of resistance from physicians and patients. André Dekker: "Our dream is to apply the Personal Health Train in real time. For example: A physician could be in a room with a patient and use the patient data and the predictive calculation model to provide an indication of the outcomes that could be expected for this person in case of different treatments. And this could be used as the basis for making a choice. Which would enable us to make decisions today based on data available today." Currently, it can take years before the insights from the treatment practice are available to support the decision making. Working with real time data was tested in a trial with partners in Maastricht and Rome. A prototype was made on a project basis, but there are no incentives in the system to expand this.

Due to the success of the PHT with cancer data, MAASTRO is expanding to other fields of expertise. They have started expanding internationally to include cardiovascular diseases and metabolic syndrome. On a national level, they are expanding to include data that is not recorded in the hospital, such as socio-economic data of patients or the costs of a particular treatment. This knowledge can be obtained by collaborating with, for example, Statistics Netherlands (CBS) for address data and Vektis for cost-effectiveness studies. By using this data, it is possible to gain insights that are currently difficult to acquire. These new insights can contribute to policy advice on, among other things, the effectiveness and costs of certain treatments. Which is why MAASTRO is also working on linking data that hospitals do not have, but which are important for the public health. This may also be data that citizens have recorded themselves.

"People often think that the Personal Health Train requires a lot of resources to be developed, but that is not the case. A lot of elements are available as open source, which means that what is needed most is manpower"

Challenges for the Personal Health Train

The most difficult thing about the PHT is to open up data using the FAIR principles. It is difficult because there are several Electronic Patient Dossiers (EPDs) and they work in different ways. Making the data captured in EPDs accessible and FAIR takes a lot of time and effort. What makes it even more difficult is that a lot of data is unstructured when it is captured, making it even more difficult to apply the FAIR principles. In an ideal situation, MAASTRO would prefer not to use EPDs, because they often are not the original source: An EPD often summarises many different applications used at a hospital. The data from those applications should be directly accessible to the PHT. This is better for the quality of the data and the outcome of the analysis. MAASTRO is developing tooling to facilitate the access to EPD data. MAASTRO also uses text mining and language processing. At the moment, it takes 3 months to start a study in a hospital because the data is not easy to access. That is why MAASTRO is working on technical solutions that facilitate this, for example through automatic data discovery and FAIR principle application.

One important point MAASTRO notes about working with the PHT is the imbalance in data economy. Hospitals need to apply the FAIR principles when making their data accessible. This takes a lot of effort, time and money and also involves risks. If the hospital data is leaked, the GDPR will impose a fine. Hospitals run all the risks and incur all the costs, while the users of these data have all the benefits. According to MAASTRO, this imbalance in the data economy is a fundamental problem.

Plans for 2020

In 2020, MAASTRO will focus on doing research and developing new things. The aim is to extract the PHT from the study and to transpose it to health care and registrations. The

focus is on scaling up, enabling citizens and policy makers to also make use of it. MAASTRO also wants commercial parties to support and implement the PHT in health care. According to Dekker, a number of steps are required to scale up the PHT to a national level. Firstly, there should be more awareness about the PHT among potential users. Secondly, the PHT should be withdrawn from the research sphere. In collaboration with public parties, the business world can improve its professionalisation by offering the PHT or creating a FAIR data station. The role of parties with a public mandate is to define the frameworks and basic principles for the PHT implementation. The aim is to provide the PHT as a service. A third necessity is that there should be some good examples of implementation in health care. These should demonstrate the added value of the PHT. For example, how registration pressure can be reduced, or how to transfer information from a hospital to a patient. These sample implementations show that the PHT can be used for health care and health policy.

Links

Maastro
PHT Network

5 Medical Data Works

Highlights

- Medical Data Works has been working on the Personal Health Train since 2019
- It is particularly interested in the PHT's healthcare facilities infrastructure and has started project 'Railway'.
- Medical Data Works is working on software for the development of the PHT rails, based on a prototype by MAASTRO.
- Based on this prototype, IKNL is also working on the further development for an infrastructure.
- For MDW, a common infrastructure is important, enabling the various Personal Health Train providers to connect easily.

Tim Lustberg

Chief Executive Officer

Medical Data Works

Tim Lustberg works at Medical Data Works[7]. Medical Data Works is a spin-off company of MAASTRO Clinic that deals exclusively with technical innovations such as FAIR data and the Personal Health Train (PHT). It has been working on the PHT since November 2019. It is currently still in the exploratory phase. Medical Data Works collaborates with MAASTRO and MRDM. It would like to use the PHT for personalised medicine.

The Personal Health Train possibilities

Medical Data Works is interested in using the PHT, because it offers the ability to combine data where it is currently very difficult. The problem is that a researcher or a company does not have data, but a hospital does. Currently, it takes a lot of effort to access and share that data. To properly unlock data, Medical Data Works wants to apply the FAIR Data principles. According to Tim Lustberg, FAIR is important for the use of the PHT. He also believes that the PHT is technically a very nice solution. This technical solution means that data remains at the source; the analysis is brought to the data; and this data can be used for various purposes, such as personalised medicine and research.
Medical Data Works has started to explore the market and is writing a business plan.
In addition, an initial project will start with Zorginstituut Nederland, where they will explore whether the PHT technology can calculate important indicators for national registration.

[7] Tim left Medical Data Works in March 2020.

Deployability of the Personal Health Train

Medical Data Works wants to use the PHT for research and personalised medicine. It is particularly interested in the PHT's healthcare facilities infrastructure and has started project Railway. Medical Data Works is working on software to develop the rails for the PHT. This software is based on a MAASTRO prototype. Medical Data Works hopes to develop, with its partners, the analyses (trains) that will be carried on the rails. Furthermore, according to Tim Lustberg, the possibility of personalised medicine would be a tremendous development. "If, for example, a physician inputs patient data and immediately receives a response from the system about similar patients from around the world, that would be great, of course." For him, it is important to create a community around the PHT, where everyone can ask questions. And where there is clarity about agreements regarding the communication and interpretation of the PHT.

"If, for example, a physician inputs patient data and immediately receives a response from the system about similar patients from around the world, that would be great, of course."

Challenges for the Personal Health Train

What Tim Lustberg has noticed in the development of the PHT is that hospitals are sometimes reluctant to make their data available for the PHT. According to Lustberg, this could be because there is no platform yet where data providers and users of the PHT can come together to build trust.

Another point he is facing is that there are several Personal Health Train implementations that are not mutually compatible. To solve this problem, Tim Lustberg would like to see an ecosystem where people are rewarded for delivering data, as well as a shared rail network, because this is the only way to use data from the system. "To keep the train metaphor going: It does not matter whether the 'train' is provided by British Rail or Eurostar, they go on the same track. A common infrastructure that enables the various rails (Personal Health Train providers) to connect easily is in everyone's interest. And that requires a national or international standard. This standard should come from a community, where users with questions and/or data can easily find each other."

The question around PHT that Medical Data Works still wants answered is: When is informed consent required to process data? In particular, with regards to the legal and ethical issues surrounding the PHT, we will have to look carefully at what is possible. The PHT is a nice concept. However, the challenge (or even the risk) of making data available is that others can also use these data in a negative way (for example, for benchmarking). The PHT can be used to discover which healthcare providers are performing well and which are less successful. These are sensitive insights that should not be available to everyone because this information can be used in a negative way. How to deal with these challenges and figure out the best way forward is also an important issue.

Plans for 2020

In 2020, Medical Data Works wants to invest the revenues from the consultancy in a business plan for developing the PHT trains (algorithms and analyses). Medical Data Works also wants to focus on three points. First of all, the feasibility of the PHT as a business-level concept. Secondly, it will focus on setting up a platform for people working with the PHT. Finally, Medical Data Works is committed to applying for subsidies to create a business case to develop the trains.

Links

Medical Data Works
Railway project

6 Netherlands Heart Registration (NHR)

Highlights

- The NHR has been working on the Personal Health Train since 2019
- The NHR would like to use the Personal Health Train to get answers to clinically relevant questions that use data from other registration organisations

Saskia Houterman

Epidemiologist

Dutch Heart Registration (NHR)

The Netherlands Heart Registration (NHR) is a quality and registration organisation that monitors and promotes the quality of care in the fields of cardiology and cardiovascular surgery. The NHR is acting on behalf of mandated physicians, who take part in registration committees on behalf of the hospitals, and who jointly give the NHR terms of reference for processing personal data. The privacy committee is responsible for legal issues on behalf of the hospitals. The NHR facilitates the registration committees. In 2019, the NHR introduced the Personal Health Train (PHT), and started a practice trial the same year. If this practice trial completes successfully and delivers promising results, the NHR is interested in a more in-depth study of the PHT. The NHR would like to use the PHT to get answers to clinically relevant questions using data from other registration organisations.

The Personal Health Train as a solution

For the NHR, the PHT may be a great solution to keep data at the source, while performing analyses in conjunction with data from other sources. The NHR is interested in the PHT because it allows them to link data to other (registration) data, which enables them to obtain more insights and answers for quality improvement. At the moment, it is difficult to combine data from different registrations, because the laws and regulations in this area are often restrictive. Sharing or combining data with other organisations is not permitted. The PHT can enable this and offers a nice solution. Keeping data at the source eliminates the need to send sensitive (patient) information. It does allow analyses to be performed without the data leaving a trusted environment, providing only the answer to a (research) question. The NHR is interested in using the PHT because data from other organisations can also be reused for other purposes. Saskia Houterman: "I think the workload will decrease if it becomes possible to use data that other registrations have already collected. This data can be used by linking it, so it takes much less money and time for you as an organisation to use it. It also prevents duplicate registration at hospitals." To make the most of the PHT, the FAIR principles are an important starting point for putting it into use.

Practice trial of the Personal Health Train

The NHR was introduced to the PHT in early 2019 and started a practical test in collaboration with the Dutch Institute for Clinical Auditing (DICA) and the Zorginstituut Nederland (National Health Care Institute). The practice trial is examining whether breast cancer patients who received radiotherapy were more likely to undergo a percutaneous angioplasty than breast cancer patients who had not received radiotherapy. To answer this question, two data sets, one managed by DICA and one at the NHR, are being analysed by the Personal Health Train; the answer to the question is returned to the researcher. The main purpose of the practical trial is to test whether the concept of the PHT is realistically feasible within vertically partitioned studies. This practical trial is expected to be completed in 2020.

> "I think the workload will decrease if it becomes possible to use data that other registrations have already collected"

Deployability of the Personal Health Train

The NHR wants to use the PHT to enable the registration committees and scientific associations NVVC and NVT to receive answers to clinically relevant questions, which can improve the quality of care for cardiac patients.
These clinical questions are raised by physicians and patients working at organisations that are affiliated with the NHR. Houterman: "So these are all the questions you can answer by linking multiple registrations – this can be revealing and provide insight to improve patient care. This will, in turn, benefit the patients. And that can be applied on a very wide scale because you often don't need the data; you just want an answer to your question." In the end, it would be a nice use of the PHT if international data could also be used to answer questions and gain new insights.

Development challenges of the Personal Health Train

The biggest challenge to make the PHT succeed in practice is to find out what is legally possible when analysing data. That is also the biggest challenge in the practical trial at the moment. Saskia Houterman: "I think that many registrations are faced with these issues. Registrations have made certain agreements with the hospitals that supply the data. Those hospitals still own the data and give the NHR terms of reference in the registration committees. We at NHR cannot do anything with the data outside the scope of our service agreement with the hospitals. This agreement explains what we can and cannot do with the data. So how we can legally apply the PHT within the framework of that service agreement is still a challenge."

The experience gained with the PHT made the NHR realise that how well they deal with the semantics of the data when data is linked, is a very important point. It is important to consider whether the underlying definitions and principles are consistent when talking about the same things. Another question the NHR asks is: How can the privacy of (patient) data be guaranteed? The NHR has made agreements and conditions for how data can be used. These must be taken into account.

Developments in 2020

As soon as the practical trial is completed, the NHR will collect the experiences and outcomes from the practical trial. Based on this information, the NHR will identify points of interest and pitfalls so that an appropriate solution can be found. The outcomes of the practical trial will determine the NHR follow-up. If the outcome of the practical trial is positive, the NHR will focus on the actual implementation.

Links
NHR Website

7 Netherlands Organisation for Applied Scientific Research (TNO)

Highlights

- TNO has been working on the Personal Health Train since 2017
- TNO has developed several applications that use multi-party computation, federated learning, distributed learning and self-sovereign identity (SSI).
- TNO is developing these technical components in order to eventually implement the Personal Health Train.
- TNO wants to use real life data (wearables and lifestyle data) in health care and prevention.

Carla Rombouts-Gordijn
Senior business manager
TNO

Carla Rombouts-Gordijn works at the Netherlands Organisation for Applied Scientific Research (TNO), an independent research organisation focused on making knowledge applicable for organisations and governments. TNO's interest in the Personal Health Train (PHT) started gaining pace two years ago, when TNO expressed its ambition to apply real life data (wearables and lifestyle data) in health care. TNO's ambition is to use the PHT for prevention and thus focus on healthy citizens. TNO is collaborating with other organisations in a Public Private Partnership called Connect2healthconsumer.

Prevention and the Personal Health Train

For Rombouts-Gordijn, the PHT offers added value when it is used to apply real life data for prevention. Health care is still very focused on curing sick people, whereas it should be focusing on disease prevention to achieve a lot of health benefits. Many solutions can lie in prevention, especially for the increasing number of lifestyle-related disorders. Prevention plays a small role in current health care, partly because healthcare professionals can gain only limited insight into people's lifestyle and the necessary information is not available. However, the PHT and Personal Health environments could change all this. Citizens are collecting more and more health data themselves, for example through their activity tracker or wearables. These wearables contain information about heart rate, sleep rhythm, weight, etc. This information is located in different applications. An overview of these different data collections can be compiled in a Personal Health Environment. This overview gives citizens more insight into their health. The added value of the PHT is that the Personal Health Environment overview can be made available to healthcare professionals. By making the wearables information available to healthcare

professionals, the healthcare professional and the patient or client can focus more on disease prevention. This application of the PHT also has a major social interest, as the prevention of lifestyle-related diseases also reduces the burden on health care.

Personal Health Train developments

TNO is focusing on four developments. The first development deals with the technical content of the PHT. TNO has developed several applications that use multi-party computation, federated learning, distributed learning and self-sovereign identity (SSI). TNO is developing these technical components in order to eventually implement the PHT. Secondly, TNO is looking at which health sources are relevant to access for use with the PHT. Thirdly, TNO is trying to determine how it can apply the FAIR principles to unlock its data. The purpose of this is to enable its data to be made available for use by other parties. As a last development, TNO is in discussion with several parties to start a programme that actually enables the data to be made available for the PHT.

> "The focus is now mainly on the healthcare market, but most health gains and benefits can be achieved by including real world data. That enables you to achieve real results in the field of prevention"

TNO has been researching the PHT by investing in knowledge about the concept, learning what external parties are doing, and determining what the PHT can do for healthcare professionals. Because TNO provides broad access to healthcare data in a way that guarantees the privacy of citizens, healthcare professionals are able to provide (preventive) personalised care. This is important for TNO. The TNO Lifestyle4health programme specialises in various interventions aimed at the (still) healthy citizen. TNO wants to share and develop that knowledge, and that is only possible if real world data is available.

Learning points for the Personal Health Train

Carla Rombouts-Gordijn notes that the concept of the PHT is difficult to explain to people who have less knowledge about the technical background, because it is complex. Because of this complexity, the healthcare field is also slow to respond to this development. Rombouts-Gordijn especially notes that on the policy side, it is more difficult to convince people and get them involved in this development. To clarify the advantages and added value of the PHT, we need better awareness. This can be achieved through more communication about the PHT.

What is also remarkable, according to Rombouts-Gordijn, is that the PHT is still very research-oriented. The PHT is already well-developed and has even been implemented in the scientific field. In the healthcare field, this development is lagging. The PHT has a major social interest because it can address health problems better than current initiatives can. To support the healthcare field in this development, the government should play a greater role. If the government invests in the PHT and promotes this topic more in its policy, this should motivate the healthcare field to pick up the pace.

One challenge in the implementation of the PHT is to combine wearables with data recorded by physicians. Rombouts-Gordijn: "Opening up data recorded by physicians should be achievable. However, unlocking data on nutrition, sleep, activity, etc. and integrating it into a health record and combining it with the PHT is still a big step." To achieve the use of wearables in the PHT, TNO has established a Public Private Partnership called Connect2healthconsumer. Connect2healthconsumer includes organisations that are looking for safe data-sharing collaborations to jointly invest in new technology and standards, and then share this information with each other. The goal of Connect2healthconsumer is to establish standards to enable the exchangeability of data, both technically and in terms of content. As soon as these standards are in place, the wearables information will be available in Personal Health environments and can support the care process. This also requires close collaboration between providers of Personal

Health environments and the PHT network.

Developments in 2020

In 2020, TNO will continue the development of Connect2healthconsumer. The aim is to work together to develop the technical content of the PHT, enabling the creation of a broader basis in the field. Standards can then be developed in collaboration. Within the coalition, use cases are being developed, which will focus on the Personal Health Train implementation in the prevention environment.

Links

Lifestyle4health

8 University Medical Centre Utrecht, Department of Psychiatry

Highlights

- Have been working on the Personal Health Train since 2017
- Are active in the CoViDa consortium
- Using the Personal Health Train to replicate and validate knowledge and algorithms in a different context
- Have developed a forecasting model for predicting an aggression incident at the GGZ (mental health care service), thanks to the Personal Health Train
- Two algorithms are currently being developed.
- One algorithm to predict the effect of antidepressants and the other algorithm to predict which patients will be sectioned

Floortje Scheepers

Professor of Innovation in Mental Health Care

UMC Utrecht, department of psychiatrie

Floortje Scheepers works at the Department of Psychiatry of the University Medical Centre Utrecht (UMCS). The psychiatry department has been interested in the concept of the Personal Health Train (PHT) since 2017. This interest has come from the ambition to start working in a data-driven manner. An infrastructure has been set up for this purpose. After developing algorithms within the UMC, the idea arose to build a model where knowledge and algorithms can be replicated and validated in a different context. This is called the 'Computing Visits Data model' (CoViDa). By using algorithms and models in a different context, they can be improved. The PHT is used to share algorithms and models with other organisations, without sharing data. This enables these organisations to use the algorithms with their own data. The parties involved are: the Department of Psychiatry of Utrecht University Medical Centre, the Department of Psychiatry of the Antonius Hospital, mental health care service Parnassia in the Hague and Rotterdam, and mental health care service Eindhoven (GGzE). They form a network of organisations called the CoViDa consortium.

Data-driven work with the Personal Health Train

The PHT concept is that the data goes to the analysis. For the department of psychiatry at the University Medical Centre Utrecht, the PHT means that algorithms, the knowledge gained about the development of the algorithms, and the output are shared with other organisations, without sharing the actual data. The data is privacy-sensitive because the material often consists of texts about patients, which make anonymisation difficult, and data-sharing is not desirable. By sharing algorithms, knowledge, and output with the network of organisations and by testing in another context with other patient data, the algorithms can be customised and improved based on the specific context and experience of other organisations.

The department of psychiatry decided to start working with the PHT, so that they could operate in a more data-driven way and thus make better use of practical data to improve the quality of care. To achieve this, the department of psychiatry started to develop algorithms that allow computers to analyse data. When developing and using the algorithms, the department of psychiatry quickly discovered that it was unknown whether an algorithm could be used for the entire mental health care service (GGZ), or whether it could only be used within this UMC. That question led to the ambition to create a network where algorithms can be shared in order to replicate and validate them in a different context. Based on these experiences and findings, the algorithm can then be optimised. This is the second reason why the PHT is interesting: to share algorithms in a network and optimise them.

The use of the Personal Health Train

One example of a major application of such an algorithm in mental health care is being able to predict aggressive behaviour in clients. This aggression forecasting model can better predict an aggression incident than the employees involved can with a questionnaire. These outcomes can provide valuable support to the employees by making the work and the situation safer for clients. The forecasting model performs this analysis by combining the snapshots captured by all the employees involved. Because a client has to deal with many employees and these employees all add an update to the client's file, there is no one who has the complete picture. A client's file is built up from individual stories. The algorithm can analyse these stories and make an aggression forecast based on this information that is accurate and reliable. The texts are analysed by computer applications (using text mining and deep learning). The PHT allows these algorithms to be shared with a network of organisations. The algorithms can be used to improve health care and safety.

> "The Netherlands is suffering from what I call 'pilotitis'. So many pilots, so many trials, so many experimental labs. There is little collaboration and each party is reinventing the wheel, which means that in some cases, systems are no longer compatible."

Personal Health Train developments

The organisations of the CoViDa consortium are in the research phase of the PHT. Organisations regularly exchange algorithms to validate, replicate, and improve. However, the algorithms are not yet used for clinical situations. Currently, there is one algorithm that is being implemented in daily practice: a forecasting model for predicting an aggression incident. Furthermore, two algorithms are under development. One algorithm to predict the effect of antidepressants and the other algorithm to predict which patients will be sanctioned. These algorithms are then shared within the CoViDa consortium and applied in practice so that they can be replicated and validated. To achieve the current data-driven practice and cooperation, the Department of Psychiatry of the University Medical Centre Utrecht has invested in data scientists, infrastructures and a data-readiness check. A data-readiness test checks whether an organisation is ready to work in a data-driven way. This readiness check looks at the infrastructure, what knowledge is available and whether a data scientist is available. In short, it looks at what needs to be done to make data available in a good way to start data-driven working. The UMC has also invested in the CoViDa consortium through administrative contacts, putting data scientists in touch with each other and building a platform where knowledge can be exchanged. In the next step, the organisation is involved in dealing with clinical questions. By organising focus groups and stand-ups with professionals and patients, real-world questions are collected so they can be answered with data.

The use of the Personal Health Train

Floortje Scheepers wants to use the PHT to improve health care and, more specifically,

integrate clinical practice and science. Scheepers: "I would like to see that data scientists looking at more and more real-world data; they will provide meaning to this data with professionals and patients; adapt policies, monitor their impact and learn from it; and thus dynamically and agilely further improve treatment and policies." The ambition of the department of psychiatry at the University Medical Centre Utrecht is to nationally implement this way of working. The ambition is to create a national mental network that includes different mental health care institutions and can work dynamically and agilely. And thus be able to constantly learn from real-world data, share algorithms, replicate and validate, and improve practice.

Challenges of the Personal Health Train

For Scheepers, the biggest challenge is the implementation of these algorithms. The spread of knowledge in a network of organisations is already happening. However, the most important thing is that the algorithms and knowledge will be used by professionals. According to Scheepers, it will take time and energy to show professionals the added value of these algorithms and the PHT.

Furthermore, Electronic Patient Dossiers (EPDs) in psychiatry are often challenging. Each EPD is different and uses different terms and meanings. In addition, indicators for symptoms of mental illness are difficult to translate into objective measures. This makes it a challenge to find out how to quickly apply an algorithm to another EPD. In mental health care, a lot of data is stored in text in EPDs, which is why text analysis is very popular. This also has advantages: applying algorithms to text is easier than to individual health data in an EPD, because each EPD has different structures and standards, while text is simply text. The best solution would be to develop an infrastructure that facilitates the reuse of data. Floortje Scheepers: "What you want is to work in a 'sandbox environment', not a constricting data warehouse in objective sizes. However, data is ideally suited for easy and flexible combination in an environment where you can use different tooling." Another challenge that Scheepers emphasizes is the establishment of a national infrastructure.
This is due, among other things, to the fact that the ICT structures of different organisations can differ greatly. For organisations that are less developed in the ICT field, it takes a lot of time and energy to prepare for exchanging and using algorithms. This takes time, attention and investments that everyone must be prepared to make.

Collaboration with other parties

The parties involved in the CoViDa consortium are the Department of Psychiatry of the Utrecht University Medical Centre, the Department of Psychiatry of the Antonius Hospital, the Parnassia mental health care services in the Hague and Rotterdam, and the mental health care service Eindhoven (GGzE). The focus for next year will be to expand this consortium. More and more organisations are interested in participating. There are two reasons for this. One reason is that algorithms are shared on the CoViDa platform and are available to everyone who is participating in that consortium. A second reason is that knowledge is shared about the use (or reuse) of available algorithms. Within the consortium, the algorithms can also be tested for reliability, validity and replicability. Because algorithms can be replicated several times, an algorithm is robust and reliable and can actually improve health care.

Links

CoViDa-platform

9 Vektis

Highlights

- Vektis has been working on the Personal Health Train since 2018 and has set up its own infrastructure internally.
- Vektis wants to combine the claims data with other healthcare data in a more convenient and secure way and allow others to use the Vektis data for analysis.
- According to Vektis, a common infrastructure is needed to let future users know how to get started with the PHT.

Jack Broeren

Product Owner PHT-FAIR

Vektis

Jack Broeren works for Vektis, the organisation that provides clear care insights and smart processes within the healthcare market to make decision-making and execution more effective and efficient. Vektis uses its products to enable parties to improve the quality, accessibility and affordability of care in the Netherlands. Since 2018, Vektis has been working on the Personal Health Train (PHT). Vektis has access to all the claimed care treatments and hopes to use the PHT to generate added value for the field of healthcare on the basis of these data. The data could, for example, be used to compare the quality of similar treatments and prognoses and to select the best treatment, thereby reducing costs and improving care. Vektis collaborates with MAASTRO and IKNL on the implementation of PHT.

Getting started with the Personal Health Train

According to Broeren, the PHT is primarily an infrastructure, which provides the opportunity to get analyses at the data source. The PHT is the feature that creates the Trains and Rails that allow analyses to be shared. This is inextricably linked to FAIR data; FAIR provides clarity about data provision and recording. Vektis wants to invest mainly in opening up access to the FAIR data and making it available for analysis via the PHT infrastructure. For Vektis, it is important that the data remains at the source, so that relevant stakeholders can perform analyses based on the same data. However, the latter is not yet possible. External parties are as yet unable to run an analysis on the underlying data without the data leaving the Vektis environment. But Vektis is working on a means to facilitate this. Vektis also wants to combine the claims data with other care data in a more convenient and secure way. Vektis is committed to using the PHT to optimise the quality of care, for example through identifying trends in care and comparing treatments.

Vektis has invested in training employees, in implementing FAIR data and the PHT, and

has set up a dedicated production team to work on this. Vektis has set up a network by contacting other organisations and people who are working on the PHT, enabling them to spend time together and learn from each other. The organisation has also invested in local infrastructure, preparing for the technical implementation of the PHT and FAIR data. This infrastructure consists mainly of hardware and virtual machines.

Vektis has carried out a pilot in cooperation with the Dutch Healthcare Authority (NZa). This pilot tested whether a question can be put to another system that will then result in a logical answer being returned. This pilot was successful. As a follow-up to the pilot, Vektis is examining how large amounts of data can be made available using the FAIR method, and how a production environment can be set up. The focus for the coming period is mainly on developing the specific PHT.

Challenges of the Personal Health Train

One of the challenges facing Vektis was the collection of technical and substantive knowledge about the PHT. Vektis found that many organisations are working on the PHT, but that the information around the PHT is fragmented. This made it difficult to find information and learn from it. To bring parties closer together, Vektis is part of the PHT learning community. This enables Vektis to connect with people who have experience with the PHT, are facing the same challenges, can share solutions and make them accessible to the entire PHT network. Jack Broeren: "If we all do something, we can complement and strengthen each other. It is important to have confidence in other organisations and in the work that has been done.".

> "The nice part is that it's all new technology. The tricky part is that it's all new technology."

Another challenge of the PHT is that it has a steep learning curve. The PHT is not something that can be done 'just like that'. It requires manpower, if you want to use the PHT successfully. Broeren: "The nice part is that it's all new technology. The tricky part is that it's all new technology".

Another challenge is how to protect privacy-sensitive information. Vektis always gives high priority to the proper handling of privacy-sensitive and competitive information. That is why Vektis is still looking for a good solution for protecting privacy-sensitive information in the PHT.

The future of the Personal Health Train

In February 2020, Vektis decided to pause the further development of PHT and FAIR data. The main obstacles were related to the lack of clarity regarding the governance of privacy and competition matters. This is essential for Vektis, since data about someone's health is extremely privacy-sensitive.

Up to now, Vektis has concentrated on finding out what is required to create some FAIR datasets and on implementing the PHT in the Vektis network. The focus was mainly on what was needed to ensure that the PHT and FAIR data provide value to the users in the Vektis network. Jack Broeren: "The idea is that data remains at the source and the analysis goes to the data. One of the things that is vital in these cases is that the result of the analysis must be good." Jack points out that it is also important for him to get the performance right. The model exists and we know what we want. However, to get the model working and scale it up is another challenge.

According to Jack, the availability of infrastructure and sufficient data stations is a prerequisite for a nationwide scaling up of the PHT. This infrastructure is a product of Vantage, developed through open source by IKNL. Broeren says that it is a barrier if the parties have to build an infrastructure themselves. If it is possible to easily install and configure Vantage's infrastructure or other tooling in a private network, that would be

appealing. And save on development costs. The focus can then be on configuring the data correctly, for example. In other words the parties could then set about making their FAIR data available. According to Jack Broeren, what is needed is a common infrastructure to let future users know how to get started with the PHT. So people can use it without difficulty. Connecting to the PHT should be very straightforward. As soon as Vektis resumes the development of FAIR data and the PHT, they will continue in that direction.

Links

Vektis

10 Zorginstituut Nederland

Highlights

- The Zorginstituut has been working on the Personal Health Train since 2018
- The Zorginstituut is involved in various Personal Health Train initiatives that look at the boundaries and preconditions of the concept
- In 2019/2020, Pels Rijcken conducted a legal analysis on behalf of the Zorginstituut to ensure the PHT is compliant with the GDPR.

Wouter Franke

Adviser

Zorginstituut Nederland

Wouter Franke works as a consultant at the Dutch National Health Care Institute (Zorginstituut Nederland), the organisation that monitors the quality, accessibility and affordability of the Dutch healthcare system. Exchange of information plays an important role in health care, as well as in healthcare management. Among other things, the Zorginstituut is involved in researching new developments in information management with the aim of furthering promising applications and ultimately improving health care. The information management activities in the healthcare industry have generated an interest in the Personal Health Train (PHT). In 2017, the National Health Care Institute was introduced to this and soon after that, people began to study the PHT and FAIR data.

Personal Health Train developments

For the Zorginstituut, the PHT is an excellent concept for achieving better information in healthcare. In recent years, they observed several obstacles, such as an increased administrative burden through having to record the same information repeatedly, difficult information transfers and exchanges between care institutions and privacy issues when sharing data. The PHT is a great solution to these problems. Wouter Franke: "The PHT allows you to design your data-driven work in networks." For the Zorginstituut, the most important thing about the PHT is that the data is not taken in for analysis, but rather the analysis comes to the data. The data can therefore remain at the source, ensuring enhanced privacy for the data. Use of data directly from the source also meets the outcome goals[8] of the Information Council. In addition to the developments surrounding the PHT, the Zorginstituut also focuses on the FAIR Data principles, among other things as a basis for the PHT. The Zorginstituut is convinced that the FAIR principles add a lot of value to the amount of data collected in health care. The importance of the FAIR data principles was also confirmed by the Dutch government in the letters: The 'Nederlandse

[8] Medicatieveiligheid (medication safety), Patiënt centraal (the patient as focus), Gestandaardiseerde informatie-uitwisseling (standardised information exchange) and Eénmalig vastleggen van gegevens (one-time data recording).

digitaliseringsstrategie' (Dutch digitalisation strategy) [9], 'Data laten werken voor gezondheid' (making data work for health)[10] en 'de Nederlandse visie op datadeling tussen bedrijven' (the Dutch vision on data sharing between companies)[11].

Since 2017, the Zorginstituut has implemented several PHT projects. From October 2017 to May 2018, the Zorginstituut conducted a practical assessment of the PHT. The aim was to acquire knowledge and practical experience with the concept of the PHT in an operational healthcare environment. Due to the successful results of this practical trial, the Zorginstituut has conducted further research into the PHT. In 2019, the Zorginstituut, in collaboration with DICA and NHR, started a second practical assessment. For this more in-depth research, the Zorginstituut also looked at the legal preconditions for the PHT. At the instigation of the Zorginstituut, Pels Rijcken conducted an investigation and published an advisory report on the legal limits within which the PHT can function.
By resolving practical issues and helping to formulate the preconditions, the scaling-up of the PHT is getting closer. And lastly, the Zorginstituut is involved in various initiatives in the field to ensure that the knowledge gained can be shared and secured.

"For innovations, we look at what lies hidden beneath the surface. If you remove the first layer, what will you find? If we see opportunities for the whole field, as was the case with the PHT, we will start to support the field and determine the preconditions to put it into practice"

Personal Health Train objectives

Wouter Franke would like to use the PHT for areas where data in networks can be reused for secondary purposes. These include reuse of data from the care process for research, policy, customer choice, accountability, but also for personalised medicine.
The Zorginstituut would like to use the PHT to improve the provision of information and, as a result, the quality of care. Because the PHT can answer complex questions and link data from different organisations, it is possible to derive additional insight.

Development challenges

In 2019, the Zorginstituut dealt with obstacles and challenges that occur when considering the practical applicability and feasibility of the PHT, such as its scaling-up, but also legal and ethical issues. One of the legal issues that was raised during the practical assessment with DICA and NHR was how anonymised or pseudonymised data from different organisations can be linked together. For the use of the PHT, it is essential that this be possible. Encryption methods are the solution for linking pseudonymised data. By encrypting data sets with the same encryption, it is possible to link data sets, but it is no longer possible to trace them back to individual people. The application of the practical assessment with DICA and NHR was intended to explore the possibilities. Another issue that the Zorginstituut is dealing with, also in collaboration with other parties, is the scaling-up of the PHT. At present, there are different PHT networks, all based on mutual trust. However, this basis of mutual trust is not scalable, especially when collaborating with international countries and organisations. That is why the Zorginstituut is developing agreements and standards for a network of trust. This development will be done in collaboration with the PHT community. Scaling up the PHT is one of the biggest challenges the field faces in this development.
In addition to the above developments, Zorginstituut and the same community have recently created a joint PHT roadmap, which includes the joint vision to reach agreements about the PHT. Information about the PHT vision and developments is published on the

[9] Nederlandse Digitaliseringsstrategie, The Hague, June 2018
[10] Data laten werken voor gezondheid, The Hague, November 2018
[11] De Nederlandse visie op datadeling tussen bedrijven, The Hague, 1 February 2019

PHT website. The PHT website also provides a platform to inform interested parties about the options the concept has to offer and enable them to ask questions or participate in the community.

The focus for 2020

This year, the Zorginstituut is focusing on the scaling-up of the Personal Health Train and the challenges this presents. The Zorginstituut remains active in the PHT community to brainstorm with other parties about different issues. The Zorginstituut will also implement two practical assessments, using MAASTRO and NICE. The Zorginstituut is also working on increasing knowledge in the healthcare sector around the PHT.

Links

- Zorginstituut Nederland
- Final report practical assessment FAIR data
- Final report practical assessment Personal Health Train
- PHT website
- Advisory report Pels Rijcken

11 Association of Dutch Healthcare Insurers (ZN) and health insurance company CZ

Highlights

- Working on the Personal Health Train since 2018
- CZ has an ongoing PHT programme 'My Best Treatment' to identify the most appropriate treatment for an individual patient.
- The PHT is used to connect the algorithms that calculate this estimation to the outcome data

Monique te Velthuis

Policy adviser

Zorgverzekeraars Nederland (ZN)

Tjerk Heijmens Visser

Adviser

Health care insurance company CZ

In this interview, Monique te Velthuis of the Association of Dutch Healthcare Insurers (ZN) and Tjerk Heijmens Visser of the health insurance company CZ explain the added value and promise of the Personal Health Train (PHT) for Dutch health insurance companies. CZ has a programme that actively uses the PHT, called My Best Treatment; ZN is following developments around the PHT with great interest. ZN and CZ recognize the need for a technical infrastructure and standards for the PHT, and are therefore pleased that its development in the healthcare industry is now gathering momentum.

Collaborating on developing standards

The Association of Dutch Healthcare Insurers and CZ are active in the PHT network[12]. ZN keeps abreast of developments through this network. CZ is actively involved in the Personal Health Train learning community. For example, the Personal Health Train learning community is creating a set of standards and infrastructure choices that are needed to make the PHT work on a national scale. CZ and ZN consider it important that these standards and infrastructure are established with the collaboration and support of the healthcare field. CZ is a health insurance company that is pioneering the implementation of projects with the PHT. ZN monitors PHT developments at healthcare insurers and in the field, and brings their value to the attention of healthcare insurers and other organisations

[12] The Personal Health Train network

to create awareness. Monique te Velthuis: "I am following the developments surrounding the PHT with great interest. I expect the PHT to be of great value for quality records, for example. For quality records, a lot of data is exchanged between hospitals, data processors and quality records. That process could be improved and made more efficient, and the PHT is certainly an interesting development in this context." At the moment, ZN does not carry out any PHT projects, which is why it is in close contact with CZ, which is actively involved in this. CZ's PHT project is called My Best Treatment.

My Best Treatment

In CZ's vision[13], the topic 'personalised medicine' is based on three pillars: the discussion between physician and patient (joint decision-making), advanced diagnostics and decision-support systems. The first pillar, the discussion between physician and patient and joint decision-making, is important because a patient can only make a good decision if they can communicate with the care provider about that decision based on correct information. The second pillar, advanced diagnostics, provides new ways to run diagnostics, supported by new techniques. These can be techniques to determine what condition a person has, what the best treatment is, or what options are available to that person. These new techniques generate more data, and this data needs to be analysed again. That is when the third and last pillar comes into play: decision-support systems. These are systems that analyse the data and provide an advisory report based on the analysis. This advisory report should be formulated in such a way that it is accessible and understandable to both the healthcare professional and the patient.

"We support initiatives such as the Personal Health Train because we think it will be an important addition to the healthcare field"

The three pillars of CZ's vision come together in the My Best Treatment programme. This is a CZ project in which the Personal Health Train has an important role. The goal of My Best Treatment is to take unnecessary and/or superfluous oncology treatments off the market. Currently, 70-80% of cancer patients receive treatment that does not produce the desired effect. The aim of My Best Treatment is to halve the number of unnecessary treatments and associated adverse effects over the next 10 years, by identifying the most appropriate treatment for an individual patient. Personalised care is combined with maximum patient participation in the decision path[14]. One example of the application of My Best Treatment is the breast cancer selection tool. CZ is involved in the development of a personalised selection tool for women with breast cancer. This tool can show whether a patient has a high or low risk of breast cancer recurrence. Armed with the outcome, physician and patient can decide together on whether chemotherapy [15] is needed. The PHT is used to connect the algorithms that calculate this estimation to the outcome data.
To achieve the goals set for personalised medicine and My Best Treatment, CZ's experience with My Best Treatment teaches that the healthcare sector requires adjustments when it comes to culture, funding and infrastructure. In the field of infrastructure, the PHT plays an important role. Tjerk Heijmens Visser: "Realizing personalised medicine is a multi-year process, which is why we support initiatives such as the PHT because we think this can be an important component".

Personal Health Train possibilities

The PHT offers several possibilities for achieving the goals set by CZ. For example, the PHT contributes to the development of artificial intelligence in the healthcare industry by making the available data accessible to researchers. This allows for the faster and better development of decision-support systems for physicians and patients. It also means that

[13] See Visie CZ and Skipr Artikel
[14] My Best Treatment
[15] My Best Treatment examples

data can remain at the source. If data can stay at the source, it means better security and privacy. Another important element is that the insurer can provide insight into the efficacy and cost-effectiveness of healthcare innovations. Clinical studies are used to quantify this. The results of these studies determine whether a healthcare innovation can be applied and how it is funded. Determining the efficacy and cost-effectiveness of healthcare innovations is complicated. The current method of assessing healthcare innovations is often time-consuming and does not always follow an efficient process. CZ expects the PHT to improve this: by making the value of healthcare innovations sooner understood and thereby accelerating their implementation. The Personal Health Train enables the investigation of the clinical effectiveness, and the calculation (in parallel) of the efficacy and cost effectiveness. Because these studies can be carried out in parallel, this process is much faster and more efficient. CZ's vision of the future of the PHT is to use that technique to develop better personalised care.

Development challenges

The biggest challenge for My Best Treatment is the scaling-up of personalised medicine. If Personalised Medicine project has demonstrated that it works in a number of hospitals, it does not mean that it works in all hospitals. This has to do with a difference in patient population and accessibility of data. Another relevant point is the determination of the added value of these techniques. Healthcare insurers and healthcare providers will have to take a stand on when a healthcare innovation has added value. Once an innovation has real added value for health care, the healthcare providers and healthcare insurers must strive to scale up these innovations. In addition to scaling-up and determining the added value, cultural aspects in health care also play a role. Accepting the PHT and decision-support systems requires a different way of thinking. For example, how healthcare professionals and insurers can trust their data to be used properly by researchers, and that the Personal Health Train is safe. The same applies to confidence in the quality of decision-support systems.

The funding of analyses using the PHT is still an open question. Once there is an infrastructure that healthcare providers can use, analyses are also performed based on their data. The use of the infrastructures located at hospitals costs money. That raises the question of who is going to fund these costs. For example, by setting up a mechanism that will compensate healthcare providers for the use of their data for research. The same applies to the compensation of the decision-support system developers: they can be paid for using their algorithms or be paid for the outcomes for health care to which their algorithms contribute. These are issues in which ZN plays an important role. Monique te Velthuis: "I consider the role of the PHT in the evaluation of care activities to be very important. With the help of the PHT, we will be able to monitor the direction in which it is going, ensuring that decisions can be taken at any moment to avoid duplication of costs or to create technical features."

Focus 2020

This year, CZ wants to continue the My Best Treatment programme and chart its results. For example, to demonstrate the added value of the PHT through practical applications. CZ also remains active in the Personal Health Train Learning Community to help further standardisation. CZ will also focus on certain privacy aspects of new data-based techniques, ensuring that the technology can be applied in a privacy-friendly manner.

Links

My Best Treatment
The Personal Health Train network
Vision CZ

Glossary

Artificial intelligence	Artificial Intelligence is a collective term that refers to systems that automize specific human stategies for reasoning and decision-making. Artifical Intelligence may for amongst others refer to Natural Language Processing, Knowledge Representation, Machine Learning and Robotics.In the context of medical data, Artificial Intelligence mostly refers to machine learning.
Deep learning	Deep learning is part of a larger family of methods of mechanical learning. It makes use of artificial neural networks that can analyse larger amounts of data based on examples.
Encryption	Encryption is the encoding of information using a complex mathematical formula. Encryption uses a key to make information unavailable in its original form. Your encrypted data can only be restored if you use the key. This is called decryption.
Federated/distributed learning	Federated or distributed learning are techniques which being used to develop artificial intelligence (AI) - algorithms in an automated way - in other words 'to train' these algorithms - based data that are stored at different locations. This Machine learning is thus performed in a distributed or decentral manner. This technique offers the possibility to work with data without it having to be centralised.
FAIR-data principles	FAIR-data principles are guidelines for the way data is described, stored and published. The letters 'FAIR' are an acronym for: Findable, Accessible, Interoperable and Reusable
Partitioned data	Vertically partitioned data refers to data concerning different aspects of the same person. Horizontally partitioned refers to the same type of data of different peoples.
Informed consent	Describes the agreement of a person – who has received/ access to adequate information – to e.g. participate in medical studies.
(Natural)Language processing	(Natural) Language processing is a combination of statistical techniques and machine-learning techniques. This enables to find key words from unstructured texts, in order to understand a human language with a computer program.
Machine learning	Machine learning refers to the development of

	an algorithm by letting a computer perform data-analyses. In the medical context, machine learning is applied for example to let algorithms perform specific tasks based on pattern-recognition often large and historical data-sets. Machine learning may for example be applied to enhance the speed of certain tasks or to automize them. For example to perform a prediction whether a person has a certain disease.
Multi-centric learning	When multiple centres work together on a study.
Multi-party computing	Multi-party computing is a cryptographic protocol that allows different parties to jointly compute data without centralising that data. The outcome of this activity is the result of the calculation, without the parties having seen the data of the other parties.
Neural network	Neural networks are based on the functioning of the brain. They model the data using models of neurones. These will be applied to try to solve complex tasks that cannot be solved by the current computer science and AI methods.
Knowledge representation	Knowledge representation is a disciplin that focusses on presenting information in a way that computing systems can make use of this information.
Open source software	Software that has its source code published and is freely available to everyone.
Personalised medicine	Personalised medicine is a medical model with the aim to tailor medical decisions, interventions and/or products to an individual patient. Based on a person's predicted response or risk of disease.
Privacy-by-design	When designing an information system, privacy is taken into account. This remains a focal point throughout the lifecycle of the system.
Self Sovereign Identity(SSI)	SSI is a collective term for cryptographic technologies to give users control over which personal data is shared with whom, while the recipient of the personal data can quickly verify them electronically. The aim of SSI is to enable safe and efficient exchange of digital information.
Shared decision-making	Shared decision-making refers to the joint decision by a patient together with a health care professional. This requires a proper exchange of information between the two.

Text mining	Text mining is a method to recognize patterns in unstructured text. It is a form of natural language processing. It provides way to manage the growing amount of available information.

The Personal Health Train in health care

Date
July 2020

Author
This document is a collaboration between Health-RI and ZIN

This report was produced through interviews with
Leone Flikweert, Health-RI
Wiro Niessen, Erasmus Medical Centre
Gijs Geleijnse, Comprehensive Cancer Centres (IKNL)
André Dekker, Maastro Clinic, Maastricht University Medical Centre, Maastricht University
Tim Lustberg, Medical Data Works
Saskia Houterman, Dutch Heart Registration (NHR)
Carla Rombouts-Gordijn, Netherlands Organisation for Applied Scientific Research (TNO)
Floortje Scheepers, University Medical Centre Utrecht, Department of Psychiatry
Jack Broeren, Vektis
Wouter Franke, Zorginstituut Nederland
Monique te Velthuis, Association of Dutch Healthcare Insurers
Tjerk Heijmens-Visser, Healthcare insurer CZ
We would like to thank parties that contributes to this collection of stories from the work field for their time spent participating in these interviews.

More information
Zorginstituut Nederland
Willem Dudokhof 1
1112 ZA Diemen

Postbus 320
1110 AH Diemen

https://www.zorginstituutnederland.nl/
BackofficePBI@zinl.nl

Personal Health Train Community
PHT@health-ri.nl