

# Connected Patients in Light of Big Data

THE JOURNAL 2022

Diogo Neves, Henrique Martins  
Linking Patients With Data: Brain-Computer  
Interfaces and Healthcare Innovation

Eugene Fidelis Soh et al.  
Flipping Healthcare Through a Population  
Health Stack

Srdjan Babic et al.  
Data Science in Modern Healthcare

Francisco Maestre et al.  
DICOM Metadata - A Useful Resource for  
Big Data Analytics

Jörg Schwarz  
How to Utilise the Massive Amount of  
Health Data Collected by Consumers to  
Improve Health Outcomes

Betsabe Melcon et al.  
In Search of Gold in Health Data





# How to Utilise the Massive Amount of Health Data Collected by Consumers to Improve Health Outcomes

Jörg Schwarz | Senior Director for Healthcare Interoperability Solutions and Strategy | Infor | USA

This article discusses why big data is underutilised in healthcare and why it makes sense to utilise it for healthcare improvement. It also explores how to achieve data privacy and security with big healthcare data.



## Key Points

- In a value-based care environment, data has a very high value because it can be used to achieve the quadruple aim: improve patient and provider experience, improve medical outcomes, and reduce costs.
- For value-based care to work, it is important to be proactive about the health of patients, and remote patient monitoring is a great tool to identify early signs of any emerging health issues.
- Remote patient monitoring is a healthcare specific form of big data, as a high volume of data in different formats (variety) can be collected from devices installed in the patient's home or on the consumer (wearables).
- Modern APIs that are usually based on machine learned algorithms can perform this filtering for various types of diseases and, therefore, add a broad spectrum of proactive care capabilities through monitoring to an organisation.

## Introduction

The discussion about big data in healthcare always has a predictable format, especially in Europe. It laments the massive amount of data that is collected about patients, bemoans the security concerns, and evaluates the risks. At the same time, we see that in 2019 alone over 51 million units of fitness trackers were sold in the United States – and that is only 42% of the global market. That means that over 120 million fitness trackers were sold globally, including in super privacy conscious Europe. And what data do these fitness trackers collect? They collect heart rate at rest and under stress, geolocations, speed, and sometimes even blood oxygenation or blood sugar levels. My Apple watch can even run an ECG. So obviously consumers are willing to collect this data and pay for it in the name of fitness. This data is usually stored in a cloud environment (i.e., for Fitbit trackers or Apple watch), and nobody is upset about it, not even the same people who are so upset about healthcare data security and privacy. Well, I am upset about it! I am

upset that all this data is not used more productively as an integral part of a preventive healthcare system. In this article we will discuss why big data is underutilised in healthcare today and then, once we understand why it makes sense to utilise big data for healthcare improvement, we can discuss how to achieve data privacy and security with big healthcare data.

## Big Data is Multidirectional

The conventional wisdom is that healthcare providers collect a large amount of data in a multitude of different data formats about their patients, which meets the very definition of "Big Data". Big Data is defined by a large amount (volume) of different formats (variety) and by large speed (velocity). The truth is that much of this provider generated healthcare data has been unusable for machine reading because it was unstructured or semi-structured. But the progress in AI and ML over the last two decades has made it possible to interpret medical images by using feature extraction algorithms



to make sense of the information hidden inside the images. Another large pool of unstructured data are natural language texts, such as progress notes or radiology reports. These can now also be machine interpreted, for example to transcribe a conversation between a doctor and a patient and turn it into structured medical records. High velocity of data is usually machine generated data, such as data from biometric devices – but here is the problem. Traditionally patients are only hooked

back to in the security chapter.

In a value-based care world, providers have an incentive to keep patients out of the hospital, so they want to prevent health status escalations. And for that, they need data. They need data to assess the risk of a patient population, stratify risk cohorts, and devise plans to manage these cohorts according to their needs. In such an environment, clinical data is very valuable, because it is the foundation of risk assessments and

---

## Clinical data is very valuable, because it is the foundation of risk assessments and escalation prevention

---

up to devices occasionally, as in once a year for the physical. Or of course while in an intensive care unit during hospitalisation. But the goal of a modern, value-based healthcare system is to keep people out of the ER and ICU using preventive care methods. And this is where the utility of data from remote monitoring devices lies, including data from fitness trackers.

Instead of measuring my heart rate once or twice a year during the annual check-up, wouldn't it be more useful for my physician to see how my heart is doing every day, during exercise and while at rest? Wouldn't a physician want to see trends, so they could intervene if the heart rate trends upward, to pick but one example? The frustrating thing is, I am already collecting all this data, but quite frankly it goes to waste.

The thesis of this article is that we should not only consider Big Data as the data collected during the occasional interaction with healthcare professionals, but really, Big Data in healthcare should also be all the data collected by consumers about their health every single day. Next, let us explore why the data is currently NOT used.

### Fee-for-Service vs Value-Based Care – The Data Aspect

Many western economies have implemented a fee-for-service healthcare model in which healthcare providers have a very transactional relationship with their patients. They see a patient, perform a service, and charge for the service they perform. The data these providers collect is mainly for the purpose of documenting their service and substantiating their financial claims. The value of data is relatively low in such an environment because data is rarely utilised for added value purposes. That is why in countries like Germany, healthcare data exchange is still performed by yes, you read that right in the year 2022, fax. Oddly, data protection gurus seem to deem the fax transmission of data to be super secure, because they do not object to it, whereas electronic data exchange is heavily regulated. But that is another story, which we will get

escalation prevention. It turns out, it is also good for patients – who would not choose a daily regimen of atorvastatin over a trip to the ER with a myocardial infarct because of clogged arteries?

In the United States we have been on a journey towards value-based care for over 20 years. It is a slow transition, but it is gaining momentum as we have more and more data that it works. And the government and other stakeholders have also clearly understood that in a value-based care world, data exchange is essential to keep people as healthy as possible. While HIPAA, originally enacted in 1996 to allow the exchange of healthcare information in certain cases, was a first step, the Affordable Care Act of 2009 made it mandatory for hospitals to keep and exchange electronic medical records in certain standard formats so that information could flow between care providers, and it also provided access to information for patients. The recent 21st Century Cures Act from 2016 built on this and now makes it mandatory for providers and payers to share information among each other and with the patient – if a provider refuses to share data, they violate the information blocking rule. This legislative journey demonstrates how important the sharing of data has become in a value-based care world to improve clinical outcomes, patient, and provider experiences, and reduce costs at the same time (quadruple aim).

One reason why data from remote monitoring devices and fitness trackers is not widely used is that in fee-for-service markets there is simply no incentive and no reimbursement mechanism for providers to use it. But even in value-based care models, remote monitoring data is currently still under-utilised, and this has infrastructural causes.

### Big Data from Remote Patient Monitoring

Over the last decade or so we have gotten pretty good at aggregating clinical data in various Electronic Medical Record (EMR) systems. This is good during hospitalisation preventing



medical errors, for example prescribing a drug that is inappropriate for a particular patient because of their medical history or other medications, and it is also good during transitions in care, when a specialist knows the medical history of a patient that was referred to them. These EMRs also record lots of biometric data during a visit or a hospitalisation, but most consumers are most of the time not in a hospital connected to a system that feeds data into the EMR – introducing remote patient monitoring.

Remote patient monitoring truly generates big data. It is machine generated, structured data of high velocity and volume. Take for example [Vironix.ai](#), a little start-up in the United States, that has developed a remote patient monitoring application. The idea is to use inexpensive devices and publicly available APIs to collect and send a continuous stream of data to their AI platform, where noise is separated from actionable data to identify pulmonary or cardiovascular issues. Noise in this context means that most of the time data will show that everything is normal – whatever normal is for each individual patient. When the algorithm flags an abnormality, a care provider will look not at all the data, but the specific information and automated analysis that is useful for a physi-

That leaves us with two problems – how do you identify useful data, and how do you get it into the EMR? Surprisingly, there is a single answer for both problems: APIs

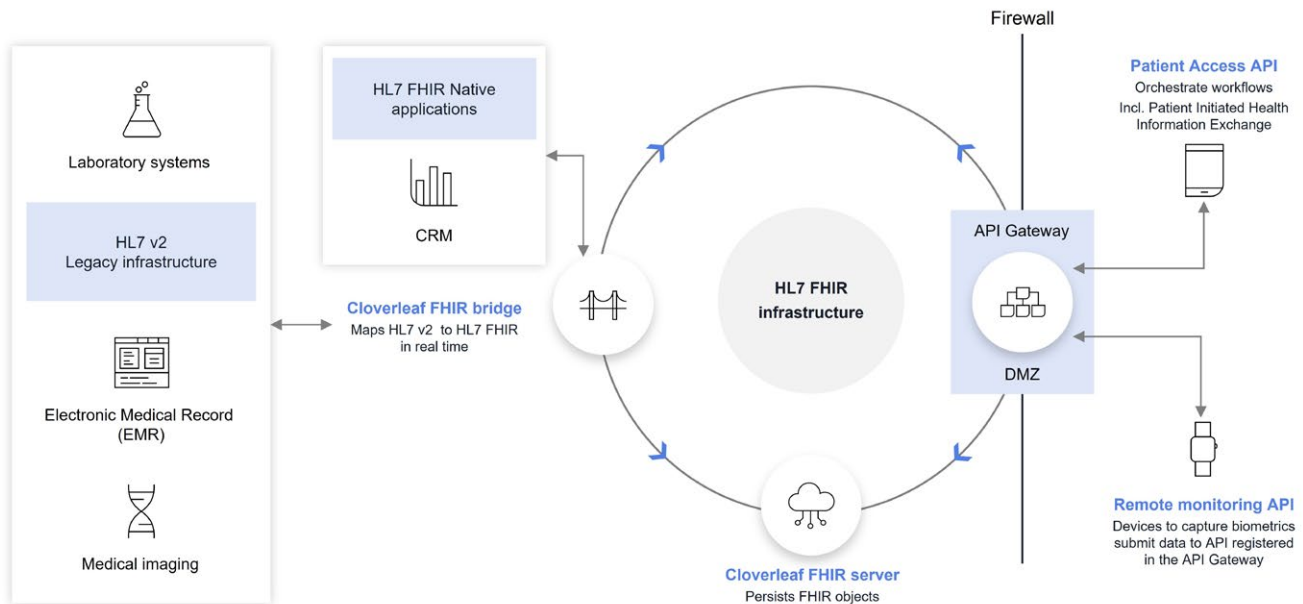
### The Modern World of Orchestrated APIs

Modern applications are built to utilise microservices through Application Programming Interfaces (API). This allows a service/application to learn new skills by interfacing with another service/application for a particular skill. A scheduling application can for example use an API to incorporate a triage service to allow a user to schedule the appropriate service by answering a few questions. The scheduling application keeps focusing on scheduling, and the triage application is focused on interpreting the user responses to find the appropriate service provider.

In the case of Vironix, the Vironix application is good at receiving and interpreting the remote monitoring data for signs of pulmonary and cardiac problems; there are other, similar applications that can interpret remote monitoring data on blood sugar fluctuations, or pacemaker data for signs of arrhythmia. All these applications need some data (for example patient demographics such as age, BMI, weight, MRN and so

Interoperability innovation delivered

## Secure Data flow for orchestrated APIs w/ legacy infrastructure



Interoperability innovation delivered | August 2022

cian. If all remote data would be funnelled unfiltered into the EMR, it would overwhelm the capacity of the care providers to interpret it and create new operational cost. If, however, useful data is filtered into the EMR, it can then be put into context with all the other data that is already in the EMR for clinical decision-making, documentation and billing.

on) to improve their filtering accuracy, and they generate data (filtering useful data from noise). The job of these applications is to identify useful data, and by using APIs, hospitals and care managers can flexibly add many different skills through an API orchestration layer. The API orchestration layer takes care of security and provides data to and receives data from those



specialised applications. It also should be able to translate the data seamlessly between different data formats. Many modern APIs require data input in HL7 FHIR format, whereas legacy healthcare infrastructure utilises HL 7 v 2.x – so the key for organisations that want to utilise big data from remote

[Bridge](#) technology with legacy HL7 v2 infrastructure (be that based on Infor Cloverleaf - the leading healthcare data integration platform in the U.S. and many European countries - or any other HL7 v2 interface engine), optimising both the flexibility of modern APIs and providing very high security.

## In value-based care models, remote monitoring data is currently still underutilised, and this has infrastructural causes

patient monitoring today is to put the right infrastructure in place to let the data flow between their existing systems and the infrastructure that interfaces with the big data world of remote monitoring devices.

### Security of Big Data in Healthcare

Fax transmission is very unsafe – I need to repeat this here because it is still the standard for data exchange in many healthcare environments. Fax data is not encrypted – anyone could spoof a phone number using VOIP and then intercept fax transmissions. Once a fax reaches its designated target, there is zero control who reads the fax or what happens to the fax. It really could not be more unsafe to transmit personal identifying health information (PHI).

In the U.S., we have HIPAA since 1996 and the HIPAA privacy and security rules since 2004 which mandate that PHI access must be limited to authorised health professionals in the pursuit of providing care for the data subject. With electronic information we can do this – data is encrypted at rest and in flight, access is controlled through password protected accounts, and any access is logged for auditing.

In the world of APIs, lots of data needs to cross the firewall and be imported into the network, and it could be possible for malicious actors to hide some compromising data inside the data streams from remote monitoring devices. These devices are, after all, out of the control of healthcare professionals and therefore vulnerable. That is why it is important to build an API orchestration layer with strong security measures, such as authenticating data sources and providing a different, shielded virtual network environment for every API.

At Infor, we built an [API Gateway](#) where each API is registered with the specific data it can retrieve and publish, and it is shielded to those specific data stores through controlled routing; at the same time, the API Gateway connects the different registered APIs through our [Infor Cloverleaf FHIR](#)

### Conclusion

In a value-based care environment data has a very high value because it can be used to achieve the quadruple aim: improve patient and provider experience, improve medical outcomes, and reduce costs. For value-based care to work it is important to be proactive about the health of patients, and remote patient monitoring is a great tool to identify early signs of any emerging health issues. Remote patient monitoring is a healthcare specific form of big data, as a high volume of data in different formats (variety) can be collected from devices installed in the patient's home or on the consumer (wearables). However, this data needs to be filtered for useful nuggets of information before the data is then fed into the existing electronic medical record infrastructure. Modern APIs that are usually based on machine learned algorithms can perform this filtering for various types of diseases and therefore add a broad spectrum of proactive care capabilities through monitoring to an organisation. Allowing massive amounts of data to flow from outside the firewall into the secure network of a healthcare provider organisation requires a secure API orchestration infrastructure, such as the Infor API Gateway. Furthermore, APIs require data from the existing legacy EMR, and they feed usefully filtered data into the EMR. For that to work, it is useful to have a FHIR to HL7 v2 real-time bi-directional translational tool such as Infor FHIR Bridge connected to the Cloverleaf API Gateway to ensure a seamless and secure flow of data from the wrist of the patient to the screen of the physician. Only when this secure pathway and intelligent filter mechanism is put into place can the vast potential of biometric data collected by patients be utilised to keep people healthy and out of the ER or ICU.

### Conflict of Interest

None. ■



**HealthManagement.org**

*Promoting Management and Leadership*