Navigating digital health

A guide to data and artificial intelligence in healthcare.

FOREWORD

Healthcare is changing. Information (data) has become more important than ever before. Artificial intelligence is starting to help organise and deliver healthcare. But what does this mean for us as individuals and how do we make decisions when these changes affect us?

We have created this guide to help answer those questions. We hope the guide will help everyone understand more about data and artificial intelligence. We want each and every person to have the confidence to make choices that are right for them about their data and its use in healthcare.

It has been a privilege to lead this project. We have had incredible support from our public contributors and subject matter experts. We have learnt a lot from everyone, and we look forward to sharing that with you.

Kelly Gleason and Jonathan Gregory

Imperial College London

June 2023

How to use this guide.

The guide has been written for you to use in any way that suits you. You can read it online, print it and read it like a booklet. There are podcasts and animated videos to enhance your learning. If time is limited, you might want to just go to the chapters that tell you what you want to know. Chapters 1, 3 and 6 would be a good place to start.

Podcast

To listen to a conversation on getting started with artificial intelligence between Stefanie Posavec, an artist who works with data visualisation, and Reshma Punjabi, a public contributor to this resource, open the camera on your phone or tablet and aim it at the QR code on the right. Centre the QR code and keep your hand steady for a couple of seconds. A link will appear on your screen, tap the link to take you to the podcast on You Tube https://www.youtube.com/watch?v=wmJDf3Ec8Lo.

CONTENT

CHAPTER 1: WHAT IS DATA? IS MY DATA SAFE?

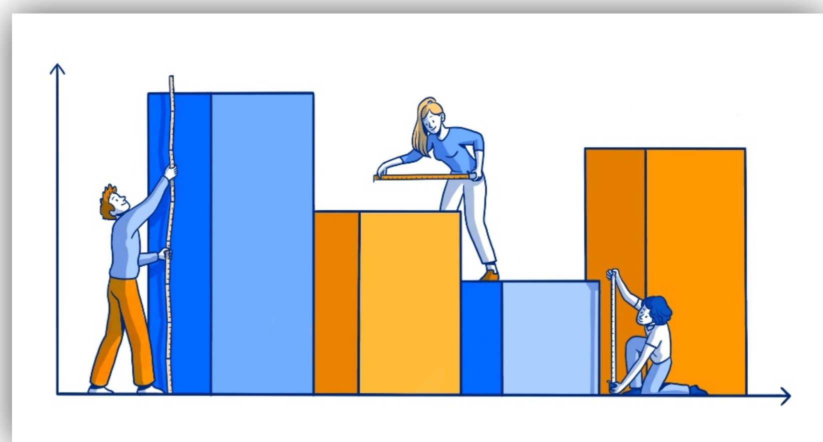
WHAT IS DATA?

In this resource we will be using ‘data’ to mean information. Data can be stored electronically on a computer, a smart phone, written in books, on a bank statement or in a photo.

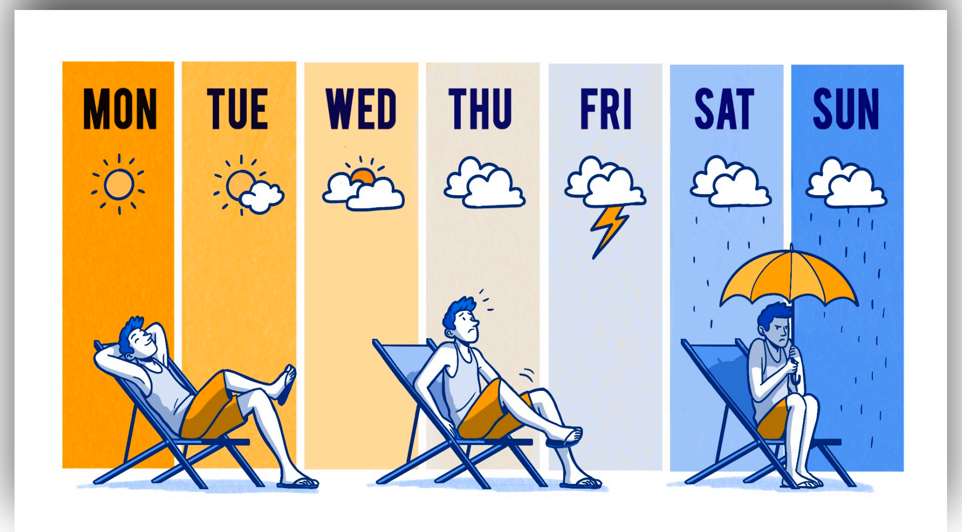
There are lots of different types of data. For example, words, numbers, dates, and times. There is a lot of hidden data in images. Some of this is words and numbers, for example the date and location a photo was taken. Also, there is data in the image itself, the colours and brightness for example.

There are lots of different ways of describing data. We will think about some of the common ways so we can all have a good understanding of all the different types of data.

Quantitative data is information that we can measure. For example, the number of children in a class or the temperature outside. It is data that tells us about a measurement.



Qualitative data is information which describes something. For example, the colours of football team kits; red, blue, red, and white stripes. The weather is also qualitative data when we talk about the weather in general such as sun, rain, snow or, here in the UK even different types of rain; drizzle, light rain, heavy rain! But not all weather data is qualitative, if we talk about the temperature or air pressure, then that would be quantitative data as it can be exactly measured.



Structured Data is data collected in a way that is easy for people to look at or for computers to analyse. If the data is in nice, neat columns or groups it is probably structured data.

Unstructured data is information which is not neatly organised. An example is written notes in medical records, or an email or letter sent between friends. There is a lot of information in these types of records. The problem is that it can be hard for people and for computers to study. We can change unstructured data into structured data, but it takes a lot of work and skill. Imagine all the written information in hospital medical records. If we read the records (and even that can be very difficult) we could find pieces of data and write them into an organised table. We could then have structured data which would be easier to use for projects and research. For example, the diagnoses and treatments that are 'hidden' in the handwritten notes could be found and typed into neat tables so they could be analysed by humans or computers.

What does all this have to do with real life?

Imagine you see your GP because you have a rash on your arm.

What would the different types of data be from your visit?

Quantitative data would be details about you or the rash that could be measured. 'Size of rash 2 cm x 5 cm , Blood pressure 120/80'.

Qualitative data would be a description of the rash. 'Itchy, red rash, worse on hot days'.

Unstructured data would be what the GP typed in your medical records. 'Saw patient in clinic today. Red itchy rash on arm for 4 weeks, worse on hot days. No insect bites. On examination it is 2 cm x 5 cm on the front of the left elbow. Looks like eczema. Try low dose steroid cream.'

Structured data would be entered into a database on the GP's computer.

Diagnosis = Eczema

Location = Left elbow

Treatment = Hydrocortisone cream

So, for all this data what would you tell someone who asked you how you got on at the doctors? You probably would not think to tell them these bits of data. You would tell them what your experience of visiting the doctors was, what the story of the visit was.

"I saw the GP about the rash on my arm. She asked me some questions, had a look and gave me some steroid cream."

Can people be identified from data?

Some people are worried that they can be identified from their healthcare data. They are concerned that private facts about their health could be shared without them knowing. These are really important concerns because we need to understand the benefits and risks of sharing our data. To help you make choices about who can access your data we are going to learn what data can be used to identify someone. We will also learn about some of the different methods that can be used to help protect our healthcare data and our identity.

Data which can be used to identify people.

One piece of data on its own cannot identify someone. It is when a few pieces of data are joined up that we start to be able to identify someone. When we think about healthcare some people have had illnesses or treatments that they do not want other people to know about. That is why keeping people's healthcare data safe is important.

Data which can be used to identify someone is called personally identifiable information. In healthcare personal identifiable information includes -

- First name and last name
- Date of birth
- Address and postcode
- Gender
- NHS number
- Phone number
- Email address

If someone with enough knowledge (and who was breaking the law) had access to 3 pieces of this information they could probably identify someone.

We have lots of other personally identifiable information which might not be in our healthcare records but can also be used to identify us. Examples include:

- Bank account details
- Passport number
- National Insurance number

There are also ‘digital’ clues that can be used to identify people for example the IP address of their device (the address of your computer, tablet or phone on the internet).

The more information we have on a person the easier it is to identify them. That means we need to only collect and use the information we really need to do the job or solve the problem we are working on.

How can our data be looked after so that we cannot be identified?

There are different ways to do this. We will describe some of them.

Encryption

This is where data is converted into a secret code so that it cannot be read. This keeps data safe where it is stored, or when it is transferred between people or organisations. It means that no one can use the data if they are not meant to have access to it. If someone illegally intercepts the data or if it is accidentally sent to the wrong place those people will not be able to read the data.

The original data is converted into a code and it is this code that is transferred. When the data is received by someone who has permission to access it they enter a password. The password 'unlocks' the data and it is decoded back into its original form so it can be used.

Encryption is used by lots of companies to protect data. It is used by the NHS to keep data secure when it is on laptop computers for example.

Anonymisation

Anonymised data is where the patient identifiable information has been removed. The data does not include items such as name, date of birth, full postcode, address, email address or phone number.

Anonymisation means that the data cannot be traced back to any particular person unless mistakes have been made in the anonymisation process.

Pseudonymisation

Pseudonymised data is data where personal identifiable information has been replaced by a unique 'code'. This has some advantages compared to anonymisation when researchers are working with data but it also has disadvantages. There is still a small risk of being identified from the data. It is very hard to identify a person from pseudonymised data but it is possible if someone has the skills and knowledge and is prepared to break the law.

Differential privacy

Differential privacy is a way of sharing data that protects the identity of the people who gave the data. The actual ‘true’ data is not shared. Parts of the real data are changed into false data. If someone sees the data, they have no way of knowing if they are looking at changed (false) or unchanged (true) data. The changes to the data are done in a way that means the overall patterns in the data stay the same. The data still represents the real world, but the true data of individual people is protected.

Encryption, anonymisation, pseudonymisation and differential privacy can all protect people's identity when done correctly. But if they are not done properly and mistakes are made it can still be possible to identify people even from anonymised data.

Some questions you might want to ask if you are asked to share your data:

- Am I being asked to share my data to support my care or treatment?
- Am I being asked to share my data for research?
- Who is collecting my data?
- How will my data be used?
- Who will be able to access my data?
- Where will my data be stored and how long will my data be kept?
- Is my data going to be used for purposes other than what I am signing up to?
- How personal is the data that I'm giving? Could I be identified from the data or will the data be anonymised?
- What is the benefit to me in giving my data? Are there benefits to anyone else (a company, the NHS, other people)?
- Will this data be freely available to be used by others?
- What will happen to my data once it has been used?

Podcast

This podcast is a conversation between Phil Booth, a data expert, and Jonathan Gregory, a doctor, about sharing your data, how to decide what to share and who to share it with. To listen, open the camera on your phone or tablet and aim it at the QR code on the right. Centre the QR code and steady your hand for a couple of seconds. A link will appear on your screen, tap the link to take you to the podcast on YouTube https://www.youtube.com/watch?v=LOjtrJMRctQ.

CHAPTER 2: TYPES OF DATA USED IN RESEARCH

Building on Chapter 1 we will now learn about data in healthcare research. Data may be joined up, shared or created to help us with research in healthcare.

IS THE DATA JOINED UP OR "LINKED" TO OTHER DATA?

Linked data involves joining different pieces of data to each other. This is really important when we need to better understand patients, their health and their care needs. Linked data can help us to have a deeper understanding of people's health.

For example, we find all the medical records of people who have had lung cancer. We then join that to information about where they live. This could be helpful. It may allow us to investigate if where you live has an effect on your chance of getting lung cancer. It may also allow us to understand if the hospitals nearby have enough lung cancer treatment services to look after the patients in that area. We may also want to link data from patient questionnaires to data about their illness or where they received care.

Is the data available to everyone?

Open data is data which is free to use (but there may be limits on what you can use it for). It is usually collected and organised by a research team.

Open data helps people with research. It can make research cheaper and it can help researchers to test ideas before they start a research project. However, like much of life, just because something is free does not make it good. Not all open data is good quality.

To avoid problems with open data, researchers need to consider where the data came from, is it accurate and does it relate to the problem we are trying to solve?

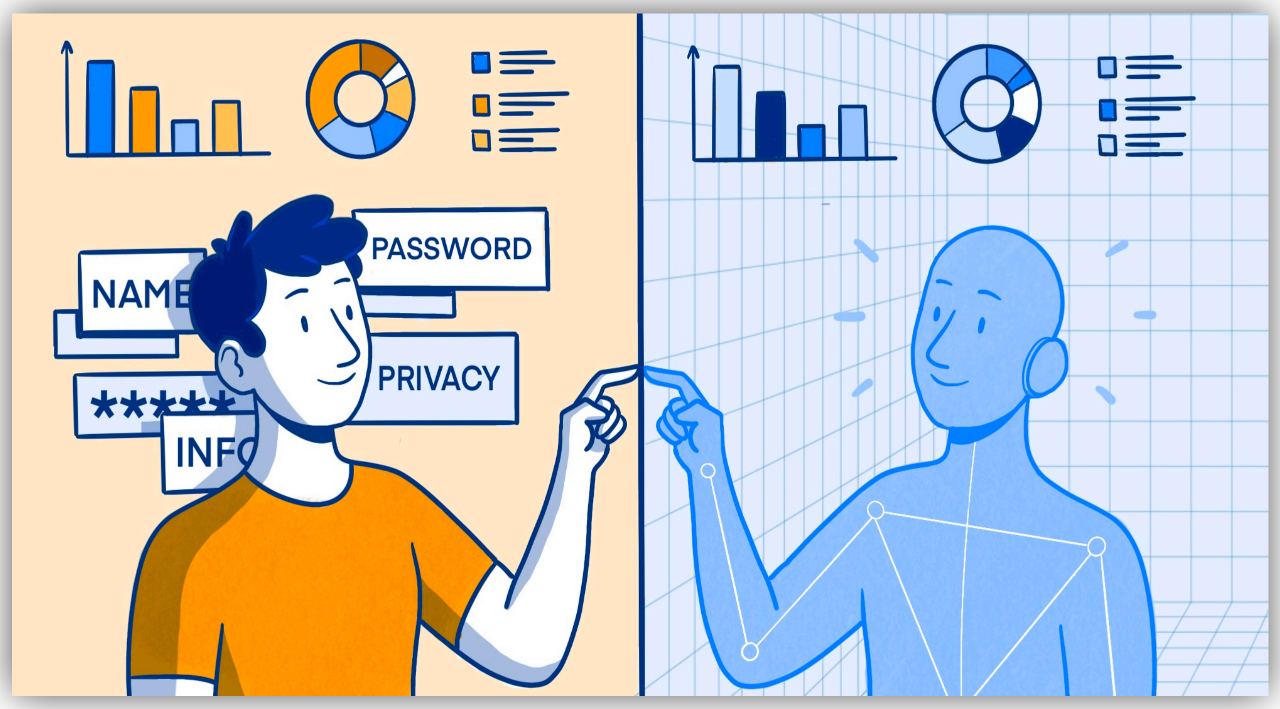
For example, there are open data sets of scans where the diagnoses are not accurate because not all of the diagnoses were made by consultants. We might still be able to use that data and learn valuable things from it but we have to be careful.

Secure Data Environments. Secure data environments are the opposite of open data. The data is held in a secure way with limited access. Usually, the data cannot be moved, it has to stay where it is and be analysed there. Often secure data environments are only worked on by people who have funding from research grants.

Open data and secure data environments have different benefits and problems so we need them both. The most important thing is that we need them to have good quality data.

Synthetic data is data that has been created to look like the real world. For example, if we had data on 1000 patients who had breast cancer and we understand lots of different facts within that data, we can create a ‘made up’ (synthetic) data set. Overall, this would have the same patterns as the real-life data but none of the ‘made up’ people would have the same data as the ‘real people' who the original data came from.

This can be helpful for sharing data securely, making sure that people’s privacy and confidentiality is protected because the ‘people’ in the data are not real there is no risk of patients being identified. But there is a different risk. If you do not really understand the real-world data, you will create synthetic data which is not a true representation of real life. That means anything you use the data for (for example research or to create a new medical device) will not work in the real world in the way you thought it would.



CHAPTER 3: PROBLEMS WITH DATA

In chapter 1 and 2 we have seen that there are lots of different types of data in the world around us. These different types of data can be found in our medical records and can be used in medical research if we give permission. In this chapter we will learn about some of the problems which can happen when working with data. In particular, we will look at bias in data; what it is and what problems can it cause.

MISSING DATA

It is easiest to think about this with an example.

High blood pressure can cause a lot of health problems in the future if it is not treated. It is important that all patients with high blood pressure have it checked regularly. A recent blood pressure reading can help doctors know how well controlled the blood pressure is and if any changes in treatment are needed.

Imagine there are two GP Practices, Practice A and practice B. Both have 100 patients who have high blood pressure. We ask the Practices to tell us about their patients who have a diagnosis of high blood pressure recorded in their medical records and when they last had their blood pressure checked.

GP Practice A sends us data for 100 patients who have a diagnosis of high blood pressure recorded in their medical records. 90 of these patients had a recent blood pressure reading but 10 patients did not. GP Practice A has "missing data" for the ten patients who have not had a recent blood pressure check.

GP practice B sends us data for 80 patients who have a diagnosis of high blood pressure recorded in their medical records. All of them have a recent blood pressure reading recorded in their medical records.

At first it seems like GP Practice B are better at checking the blood pressure for their patients because all the patients have a blood pressure check recorded. But GP Practice B has a different problem. Twenty patients have high blood pressure, but the diagnosis has not been saved in the medical records. This is also "missing data." Patients who should be on the list have been missed because the diagnosis was not recorded in their medical records.

In the real-world GP practices are very good at recording patients diagnoses and checking blood pressure results for people with high blood pressure. But hopefully this example shows you that it can be hard to compare data from different places and you have to be careful as missing data can be hard to spot – because it' missing!

Mistakes with data

There are lots of ways mistakes can be made when recording data.

For example, if we are not careful with units. Units are a standard way to measure things like height, weight, temperature. Imagine you are asked to write down your height. It asks for this in centimetres, but you write it in metres. You thought you were writing 1.6 metres; you record 1.6 cm because you didn't read the units carefully. Usually this is easy to spot as no one is 1.6 cm tall.

But it is not so easy to spot if a result is wrong for that person but could be right for someone else. You are 160 cm tall but you make a mistake, your finger slips on the keyboard, and you type 180 cm tall. It is possible that someone is 180 cm tall so it would be extremely hard to spot this mistake, but it could cause real problems for example when looking at blood results or working out the dose of a treatment according to height and weight.

Inconsistent data

This would be where we asked for height in centimetres, but some people write it in feet and inches. When data has been recorded inconsistently, it means a lot of time has to be spent correcting or cleaning the data before it can be used. It could also be where doctors have written different words to mean the same thing, for example, MRI scan, MR scan, MR all mean the same thing to doctors but can cause problems when analysing data.

DOHH WITH DATA

Researchers were investigating patients with diabetes. They collected data from lots of GP practices.

They noticed something odd about one of the cases. There was a patient that had been recorded as being dead but the patient had new entries in their medical records after the date recorded as the date the patient died.

The researchers investigated this.

In the GP database there was a section to write the date of death. This had been filled in with a date. Next to this was a box where text could be typed and in that box it had been entered 'of dog'.

The patient hadn't died at all! Someone had recorded the death of the patient's dog in the database (probably because the patient had discussed being upset after their dog died). The person filling in the records had tried to be helpful but had made the data inaccurate.

There is so much data about our lives that is easy for some of it to be recorded incorrectly, not be recorded at all or put in the wrong places on forms and in databases.

BIAS

What is bias in data?

Even when we have data that has accurate results and is well organised we can still have problems with it.

Being 'biased' is where you favour one person or group over another. We are ALL biased in some way. Usually in ways that do not hurt other people. For example football fans usually think their team is better than they actually are, most parents think their children are more talented than other people's children. But people can be biased against people and harm them either accidentally or deliberately by their views and attitudes. For example, at job interviews the interview panel will have an unconscious tendency to give jobs to people who are similar to themselves. Unfortunately, because data is affected by the world we live in, it can be biased. It can be biased by how it was collected or how it was used. That can mean that the data does not give an accurate picture of the real world or the people in it. There may be people missing from the data or too many of one group of people compared to another. Both of these mean that the data does not reflect how the world really is. We have to check that the data is representing everyone that needs to be included. We need to be aware of bias in the data and bias of the people looking at the data.

Why does bias matter?

Bias matters because it means that the data is not giving us a true picture of the world. The data is not giving us the whole story, only part of the story. If we use the data to make decisions thinking it will be good for everyone, we might find we actually cause problems for some people. We need to check if the way we collected the data or the way we have used the data has created bias. We have to be on the lookout for bias at all times. It can be hard to spot.

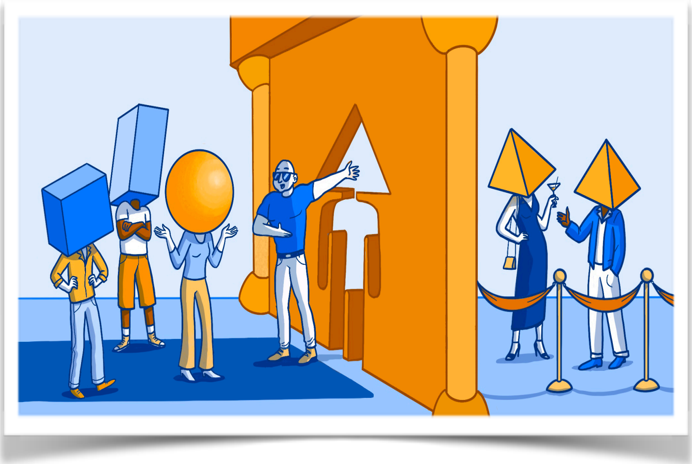
We can make mistakes that create bias in the way we collect data or in how we use it to answer questions. Here are some questions we might ask ourselves to try and make sure we have not accidentally created biased data.

- Does the data represent everyone who will be affected by the results of the data analysis?
- Have we excluded groups from the data who need to be in the data?
- Have we included everyone who needs to be in the data?
- Did we get a response rate that was appropriate across different communities?
- Could everyone understand what we were asking?
- Could everyone take part (language, time)?
- Did we ask the right questions to get accurate data?
- Have we been open minded? Have we allowed our pre-conceived ideas or thoughts to affect how the data was collected?
- Have we tested the data to see if using it could result in any unintended problems? Do the results work equally well for all groups of people?

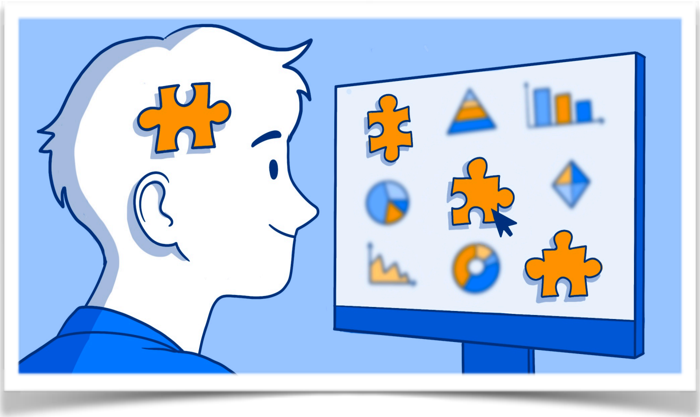
Types of bias

There are many types of bias. People who work with data need to know these so that they can be on the lookout for them. Here we talk about just a few of the types so you can see how easy it can be for bias to 'creep' into data.

Selection Bias. This is where data does not contain a balanced range of people. An example would be when we collect feedback on a hospital. Are the people who give feedback representative of all the people who use the service? We need to check that feedback comes from all groups of people. There is a risk that only people who liked the hospital or who were angry with the hospital give us feedback. That means we are not getting the opinions of people who didn't think it was very good or very bad but that might be the biggest group of people!



Confirmation bias. This is when the person looking at the data allows their own thoughts (bias) to affect how they study the data and confirm what they were already thinking. They are not studying the data with a completely open mind, and often accidentally, notice patterns in the data that agree with what they think and don't notice patterns in the data that mean the person's ideas are wrong.



Outlier bias. This is where a small part of the data has a big effect on the data overall.

Imagine you live in small block of flats. There are 9 flats. On the ground floor and first floor there are 4 flats on each floor. An estate agent thinks they are worth about £100,000 each. There is one flat on the top floor. It is as big as four flats combined. It's a penthouse with amazing views of the city. The estate agent thinks this is worth £1,000,000, 1 million pounds. Your flat is worth £100,000. You try and get some insurance for your house, and they say the average cost of flats in you block is £200,000. Why would they think that? No flats in your block are worth £200,000! The million-pound flat is an outlier, and it is affecting the average. The average doesn't give us a good idea of the real world because the outlier is having a big effect on the data.

Survivor Bias. This is where there is data missing in your sample, although you might not realise it. A famous example of this is from the second world war. The American air force wanted to improve the protection on their planes so that they were less likely to crash if they were shot. To do this they studied the planes that made it back to base to see where they had been hit. BUT that meant they were only looking at the planes that **had** made it safely back home! It was the planes that had crashed that they needed to study! They were studying the 'survivor' planes and not the planes that had crashed, which meant that the data would not give them the answer they needed, it would be biased. We have to be very careful of survivor bias in healthcare. When we look at long term data for patients this will only tell us about 'survivors'. This might not help us understand the disease in patients who have died. This can affect our results and mean that we are not learning how to make more people survive their illness.

Real world example of bias: measuring oxygen levels in blood

Pulse oximetry is a way of measuring the amount of oxygen the blood in your arteries is carrying. A clip is put over the fingertip which is connected to a machine (an oximeter). The clip uses visible and infrared light to measure the level of oxygen in the arterial blood. Often the pulse oximeter is clipped to one hand whilst a nurse or doctor checks your blood pressure on the other arm.

Oximeters were developed in the 1940’s. They were first used to assess oxygen levels in airplane pilots flying at altitude. Improvements were made to the machines in the 1970’s so that they could then be used in hospitals. To test the machines, readings from the pulse oximeter were compared to test results from blood samples which very accurately measure oxygen levels. These studies showed that the machines were less accurate when people had low levels of oxygen in the blood. Some studies also showed that pigmented skin could affected the accuracy of the machines. When the machines were used on people with darker skin, they could give a falsely higher reading than the actual level of oxygen in the blood. Both of these problems were well recognised by doctors and researchers who were experienced in using the machines. But, as the machines became more common and were used by less experienced people these problems were gradually forgotten.

Over the last 20 to 30 years pulse oximeters have been regulated as medical devices. There are legal standards they have to meet before they can be used in hospitals. But, because pulse oximeters were invented before these rules were introduced there has been a problem. When you have a brand new technology you have to meet all of the new rules with new evidence. But if you are using technology that was invented before the rules were introduced you need to show less evidence. In this case, because oximeters had existed for many years companies did not have to demonstrate lots of new evidence. Just showing they were 'as good as' old machines was enough. Some companies did test their machines from scratch. They checked how accurate they were in different skin tones and gradually improved how accurate the machines were. However, some companies did not perform a lot of retesting. They relied on providing evidence that their machines were as good as older machines. But, only being as good as these older machines meant their machines were less accurate in some patients, particularly those with darker skin. These machines are so common in hospitals that the staff using them assumed they all worked the same and treated the results the same. People forgot to make an allowance for darker skin or when people had low blood oxygen levels.

Oximeters have been designed which are accurate no matter what skin tone is. But these machines are often much more expensive. Also, these machines were not designed for patients to use in their own homes.

A group of doctors in the USA studied pulse oximetry in a large number of patients. They 'rediscovered' what had been largely forgotten. They made their rediscovery because of a large amount of high-quality data. They collected 48,097 results from 10,001 patients using one type of pulse oximeter. When they analysed this big sample, they were able to see a problem. They found that this type of pulse oximeter machine did not work as well in Black patients compared to White patients. The machines also were not as good at detecting low oxygen levels, with the machine recording a higher result than the actual oxygen level. Black patients were three times more likely to have a low oxygen level missed by the technology compared to white patients. This meant that the results could mislead doctors into thinking some Black patients were not as ill as they actually were. That will have meant some patients did not get as much treatment as they needed.

The problems with pulse oximetry teach us that we have to be much more careful to test devices with people that are the same as the people who will be treated with the device in the real world. It also shows us that good quality data needs to be collected so that we can identify problems as early as possible. It shows us that good data can help improve healthcare.

Video

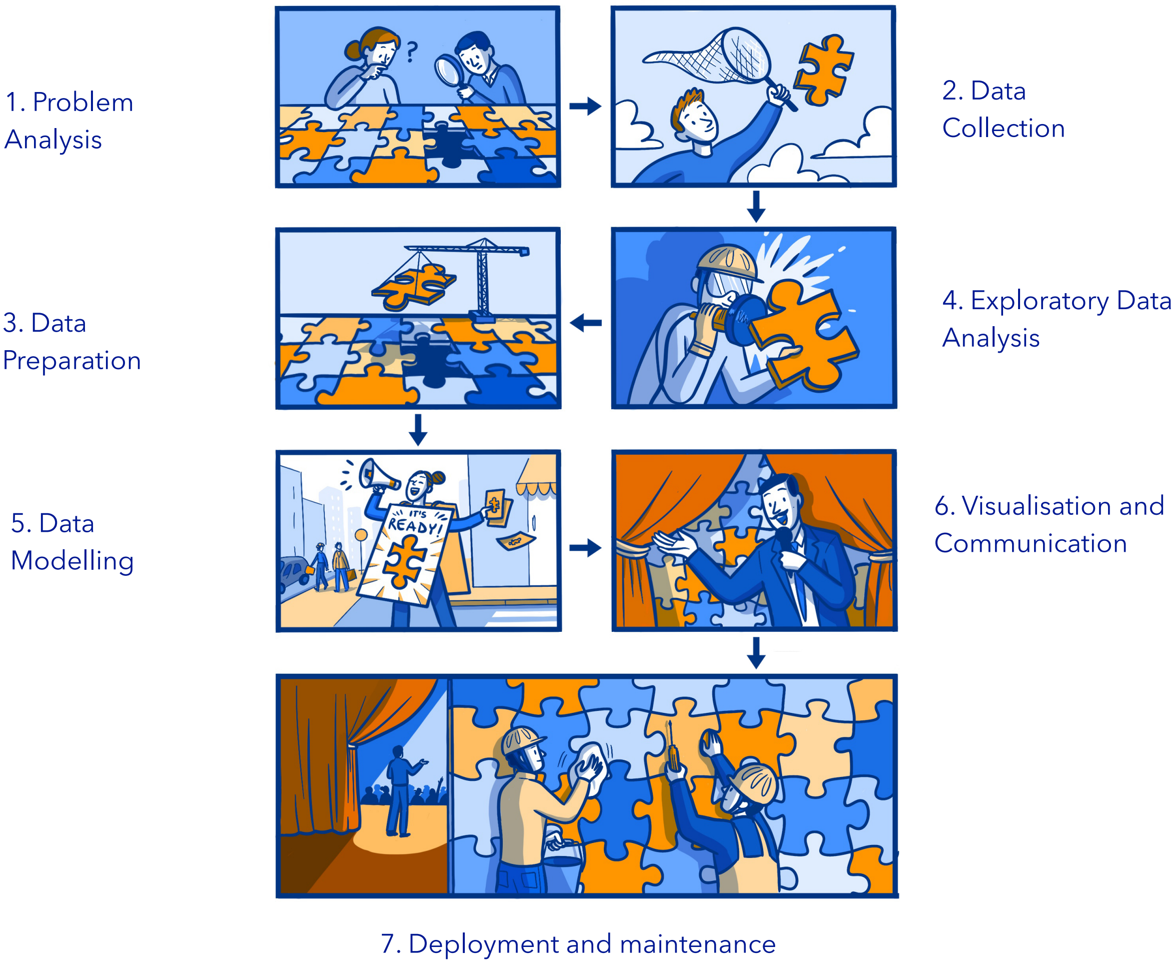
This is an animated video on bias when working with data. To watch, open the camera on your phone or tablet and aim it at the QR code on the right. Centre the QR code and steady your hand for a couple of seconds. A link will appear on your screen, tap the link to take you to the video on YouTube https://www.youtube.com/watch?v=9PQKVvjuH3A.

CHAPTER 4: A DATA PROJECT

In chapter 3 we learnt about some of the problems that can happen with data and how those problems can lead to patients being put at risk and even harmed. In this chapter we are going to learn how the data is used in projects and research. We will see the steps in a project and how we need to be careful when interpreting results.

In a factory we have quality control so that faulty items are rejected before they are packaged up and sold. We have to do the same with data. We have to check it and ensure the quality of the data is good enough before anyone uses it. In data science projects we combine problem solving, computing and mathematics in ways to help us understand data. The aim is that by investigating the data we can find and understand patterns in the data which we can use to help us solve problems. There are several stages in a data project and these show us how data is used in real life to solve problems in healthcare and other areas of the world around us.

A DATA PROJECT USUALLY HAS 7 STAGES



1 **Problem Analysis:** The first stage is to understand the problem we have to solve. The data scientist works with others so they can understand what data might be needed to solve the problem. Once the exact problem is understood then the team can decide what data they need to collect and how they might collect it.

2 **Data Collection:** The second stage is where the data is collected. It might be structured nice tidy data (Chapter 1) or it could involve something called 'scraping' where data is taken from places where it is not stored in neat tables, for example, webpages. Finding the right data may take a lot of time and effort.

3 **Data Preparation:** This stage involves two jobs. Data cleaning is the most time-consuming process as it involves 'cleaning' the data of all the errors we met in Chapter 2 for example missing and duplicate values. We also need to check if there could be bias in the data. Then a process called data transformation is used to change the data so that it is set out in the best way for analysis. It's like you getting dressed smartly for an interview or wedding, you are the same but you appear different. It's the same with transforming data: it's the same data but its appearance has been changed to make it more appropriate for the situation.

Even when data has been collected very carefully it will still need to be checked and tidied up before it can be used. Before it has been cleaned the data is called 'raw data' and it is exactly how the team will have collected it - mistakes and all. It is important that during the cleaning process we don't make changes that affect how 'truthful' the data is.

There are lots of different checks and tests done on the data at this stage. We would ask questions like these to help us check the data.

- Are there missing values?
- Are there duplicates where a person appears more than once in our data?
- Are there any outliers? If there is an outlier is this a correct result or was a mistake made when collecting data?
- Are the results all using the same measurements for example not a mixture of height in metres and feet.
- Then we need to ask ourselves, does the data seem sensible? Does it look like what we would expect? Could the data be biased?
- If the data has been collected over a period of time, was the way it was collected the same or did it change at some point? For example, if we were measuring temperature did we change the thermometer we used? If we had data on blood tests did the laboratory change the way it checked the blood during the time we have collected data?

An example of data checking

We have collected data on the weight of children in a school. We are worried that some are overweight, and some are underweight. We have results on 100 children.

✓ Are there missing values?

No, the school says there's 100 children and we have 100 results.

✓ Are there duplicates where a child appears more than once in our data?

Yes, one child has had their result collected twice and now we find out one child was off school that day. So, after cleaning this we have 99 results for 99 children and 1 child was off.

✓ Are the results all using the same measurements?

No. Most weights were in Kilograms, but some were in stones, so we change those to Kilogram in the data.

✓ Are any outliers?

Yes, one child has a weight of 350 Kg. Another has a weight of 60 kg.

We check and the 350 Kg should be 35 Kg, a zero had been added by mistake.

We check the 60 Kg as this is heavier than the other children of the same age. The result is correct. The child is very tall and muscular.

✓ Does the data seem sensible? Does it look like we would expect?

Yes. The range of weights we have collected, and the averages look like the children in the school. When we compare our data to data from other studies it is very similar.

Good news. We did a good project. Our data was good after we had cleaned it and even better the children in the school are all the right weight!

4 **Exploratory Data Analysis.** This is where we explore the data using mathematical tests to look for relationships. Is one thing related to another? Could one thing be causing another? This is a vital stage because choosing the wrong parts of the data to use will produce an inaccurate result.

The simplest tests look at the lowest and highest values and averages. More complicated tests look to see if different pieces of data are related and how they are related.

When we are exploring data, often patterns come to light. These patterns can look convincing, and we can think that one thing causes another. But there is a difference between two pieces of data being linked (correlated) and one thing changing another (causation).

This is one of the most important things in this chapter. Lots of people get this wrong – especially newspapers and that's why we get some attention grabbing headlines that don't actually fit the facts!

Correlation and causation

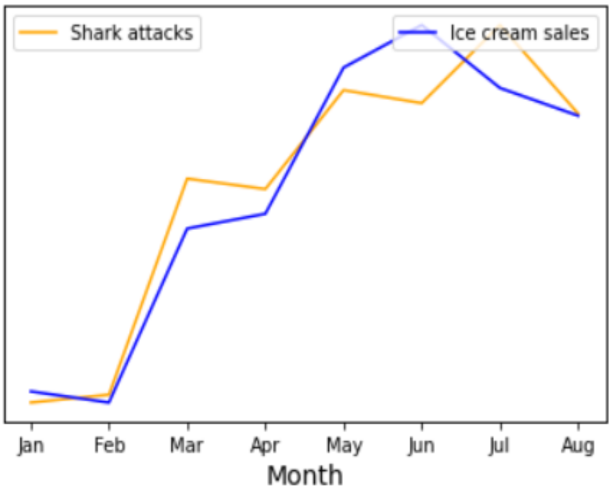
When we are exploring data, often patterns will be noticed and researchers have to work out if one thing is causing another or if they are just following the same pattern.

Imagine this headline - *Ice Cream causes people to be attacked by sharks!**

If we look at data on ice cream sales and the numbers shark attacks in the United States each year, we will find that the patterns in the data look the same. They are correlated.

But that is different from one causing the other.

Could eating an ice cream increase the chances of you being attacked by a shark? Could being attacked by a shark make you more likely to eat ice cream?



The more likely explanation is that more people eat ice cream and get in the ocean when it's warmer outside. More people in the ocean means more shark attacks. Temperature is what links ice cream sales and shark attacks. They are not directly linked to each other. Although ice cream sales and shark attacks are highly correlated, one does not cause the other.

Video

This is an animated video on causation and correlation. To watch, open the camera on your phone or tablet and aim it at the QR code on the right. Centre the QR code and steady your hand for a couple of seconds. A link will appear on your screen, tap the link to take you to the video on YouTube https://www.youtube.com/watch?v=_nG1C3jj1c8.

*Thank you to Zach Bobbitt of Statology for this example. More information on Statology in Chapter 9.

5 **Data Modelling.** The data scientist will run different mathematical tests on the data to find out which help them to best understand and explain the data. They will also see if there are tests that can use the data to make predictions. Can we find a rule from the data we could use in any situation that would give us helpful answers? For example, to make a prediction about a patient's risk of a heart attack from data on their blood test results?

The aim of stage 5 is to solve the problem we identified in stage 1.

6 **Visualisation and Communication**. We now use tools which help us to see the patterns in the data. We use graphs and diagrams to show the data in different ways. The phrase 'A picture tells a thousand words' is true when we are studying data. The words we use when we talk about data can affect what people think. Part of presenting data is being aware that the words we use and the graphs we draw affect how people interpret the data. It is easy to accidentally (or deliberately) mislead people when showing them data and the results of research.

THE WORDS WE USE WHEN WE TALK ABOUT DATA CAN AFFECT WHAT PEOPLE THINK

How we present data to readers or our audience can have a big effect on what they think. For example, we could talk about the death rate (mortality rate) from children's heart surgery being 5%. Or we can talk about the survival rate of children's heart surgery being 95%. Describing a 95% survival rate sounds a lot better than talking about a 5% death rate. This is called framing, and it can have a big impact on how we interpret data. Often the media use framing to grab your attention with headlines.

A good example of how framing can be used was in a campaign to reassure people that London was safe. The campaign proudly said that 99% of young people in London do not commit serious violent offences. Positive framing. If we wanted to paint a more worrying picture, we could frame this negatively, '1% of young people in London DO commit serious violence'.

If we wanted to worry people even more, we could use the power of big numbers. Most people don't find it very easy to imagine percentages especially when the percentage relates to a really big number. If we convert the percentage into a number, it might grab people's attention.

There are approximately 1 million people in London aged 15-25. That means 1% is 10,000 young people. Now if we really want to worry people we can say 'There are 10,000 young people in London who commit serious violence'. This feels a long way from where we started doesn't it? But it is the same as 99% of young people in London do not commit serious violent offences. Both are correct but the framing affects how we feel about the numbers.

*These examples are drawn from The Art of Statistics Learning from Data, David Spiegelhalter ISBN 978-0-241-25876-7

7 **Deployment and Maintenance.** This is the final stage. The model is tested again and then put into the real world. The model will be used to make reports and data dashboards are used to check data in real-time. The model's performance is checked and supported. This then marks the end of the data science project.

An example of a project like this would be delivery companies using data to discover the best route to take to make deliveries to speed up the process and reduce costs. Airline companies have used data science projects to help them predict flight delays and notify the passengers beforehand and plan how to catch up the delay as quickly as possible.

CHAPTER 5: THE PROBLEMS AND BENEFITS OF USING DATA IN HEALTHCARE

So far we have thought about what data is and the problems we can have when using it. We have learnt about the importance of high-quality data and how we need to be careful with bias, correlation and causation. In this chapter we will look in more depth at three examples of data being used in healthcare and healthcare research to bring together all that we have learnt so far. These examples will hopefully show you some of the problems we have been talking about. Working with healthcare data can be powerful so we all need to understand how it can affect us for good and for bad. In his book, The Art of Statistics Learning from Data, Professor Sir David Spiegelhalter shares many examples of data being used well and not so well. Here we describe two of his examples.

Problems with data in healthcare. Results can be misleading.

'Eating meat gives you cancer'

In 2015 newspapers reported that processed meat, such as bacon, ham and sausages increased your risk of bowel cancer. The way this story was presented gives us a powerful learning point.

The papers reported that experts had found that eating 50g of processed meat a day increased the risk of bowel cancer by 18%! At first this sounds like eating processed meat might be really dangerous for you.

But, digging into the data shows us that this number is right but it doesn't mean what we think it means. The headline makers tripped over the difference between absolute risk and relative risk. Newspapers struggle with this (or choose not to understand it) and it leads to a lot of misleading headlines.

"One of the biggest problems in looking at data is where people mix up absolute risk and relative risk."

Relative risk is where we compare one group of people to another. It is one group 'relative' to another. In this study they compared a group of people who ate 50g of processed meat per day to a group of people who didn't. They found that the group who ate 50g of processed meat had an 18% increase in bowel cancer compared to the other group. So far so good, it looks like the headline in the paper was correct.

BUT – before clearing all the meat products out of your fridge there is a vital question that you need to ask yourself. What was the risk of bowel cancer for the group people who didn't eat processed meat? What is the baseline risk that we all have whether we eat meat or not? This is the 'absolute' risk.

The absolute risk of developing bowel cancer (for an average person) is 6%. That means if we have 100 'average' people 6 will get bowel cancer in their lifetime, even if they do not eat processed meat.

Now we can return to our relative risk. Eating processed meat increases the baseline, the absolute risk by 18%. Increasing the absolute risk of 6 people in 100 by 18% means that 7 in 100 people will get bowel cancer in their lifetime. The 18% increase in bowel cancer is one extra person out of 100 people who ate processed meat every day.

This is still important, and people might choose to change their diet to reduce their cancer risk. It is important if we think about the numbers of people with bowel cancer in the whole country. One extra person in 100 when we think about the many millions of people in a country will mean there are a lot more cases of bowel cancer. But for one individual person making decision about their lifestyle this difference of 1 in 100 feels different to an 18% increase.

These examples show us that reporting the results of studies requires care and attention to avoid misleading people. We all need to understand a little bit about data so we can decide if we believe what we are being told.

Benefits of using data comparing hospitals.

In 1996 the General Medical Council investigated children's heart surgery at Bristol Royal Infirmary. There were concerns that the service was not as good as other hospitals. There were concerns that more children died after being treated there than at other hospitals.

Professor Sir David Spiegelhalter led a team of investigators to compare the survival rates for children who had surgery in Bristol with survival rates in other hospitals in the United Kingdom.

At first you may think this should be an easy investigation but when you are using healthcare data you have to be careful as not everything is as it first seems.

First, they had to decide what types of heart surgery they were going to focus on and how many children had received those operations. Not all operations have the same difficulty so they needed to make sure they focused on the sorts of operations most likely to be linked to problems. They also had to decide 'when' a child who died following surgery had died because of the operation. Whilst we would all agree that a child who died during the operation or a few days later died 'because of the operation' what do we think about a child that died 2 months later? What happens if they went home but then came back into hospital and died – was that death still due to the operation?

Once they had decided the answers to those questions they had to start looking at the data. There were different sources of data. There were 'hospital episode statistics', which are created by hospitals to record the work they have done but this data often contains mistakes. There was also data from national death records, and a cardiac surgical register where surgeons recorded operations they had done. When comparing all these types and sources of data, the number of operations and the number of deaths recorded were different. It took a lot of work to clean and tidy the data so that it could be used in the investigation.

The investigation predicted that, if Bristol were like other hospitals, there would have been around 32 deaths in Bristol. This number of deaths would have been expected when you consider how ill the children were and that even with good care you cannot save every child. But there had been 62 deaths over this time. 30 more than we would have expected.

The next question was were any of these extra 30 deaths avoidable? Was it that Bristol had been seeing really ill children and in fact more of them would be expected to die? These would be unavoidable deaths. Or was it that there were children who should have survived who in fact died because of the care at the hospital? These would have been avoidable deaths.

Some hospitals will have more or less deaths than the ‘average’ due to random differences in the patients who go there. For example, in one year there may be many more seriously ill patients having surgery in a particular hospital. This would make the death rate go up. There would be excess deaths compared to the average, but those deaths might not have been avoidable. It is possible for the death rate to be higher or lower than expected when in fact the hospital is doing the same as always. But, in the case of the children’s heart surgery service in Bristol the difference between the expected and observed number of deaths was large. It was so large that it told the investigators that there was a genuine problem with the service and not a problem with the data. The results of children’s heart surgery in Bristol were worse than other hospitals. The service was reorganised and restaffed.

What appeared an easy question – did more children die after heart surgery in Bristol than would be expected? - was in fact much harder to answer and that can often be the case with data.

*These examples are drawn from The Art of Statistics Learning from Data, David Spielgelhalter ISBN 978-0-241-25876-7

CHAPTER 6: AN INTRODUCTION TO ARTIFICIAL INTELLIGENCE

In the last five chapters we have been learning about data. Now we will learn about artificial intelligence (AI). We started with data because data is the building blocks of artificial intelligence. Without good data you cannot have good artificial intelligence. In this chapter we will learn what artificial intelligence is, what it can do and how it's created.

When we search the internet the search returns web pages which may be relevant. AI is what puts all those web pages into the order you see on your screen. Banks use AI to help show when bank details have been stolen. Online TV and film companies like Netflix and Amazon use AI to make suggestions for other programmes you might want to watch based on what you have watched in the past.

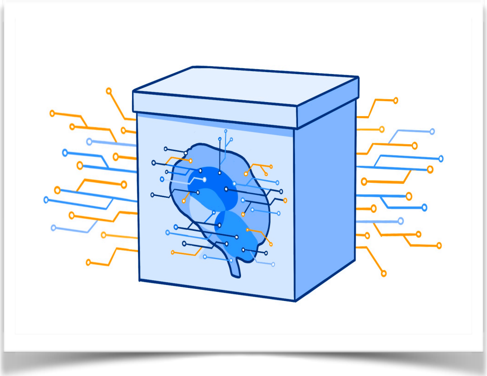
AI is not one thing; it is a range of different computer programs. If we want to explain AI in one sentence, we would say 'it is computer technology that does tasks that in the past would have needed human intelligence'. For example, translation software. In the past if we were abroad and we wanted to understand a road sign or menu we would look the words up in a phrase book. That needed human intelligence. Now there are lots of computer programs, even on mobile phones, where you can translate words. This is a task that is often now done by AI.

Podcast

This podcast is a conversation between Jonathan Gregory, a surgeon, and Danny Ruta, an expert in AI. They talk about AI and how it is created. To listen, open the camera on your phone or tablet and aim it at the QR code on the right. Centre the QR code and steady your hand for a couple of seconds. A link will appear on your screen, tap the link to take you to the podcast on YouTube https://www.youtube.com/watch?v=eXDLBsm7VVs.

How does AI work?

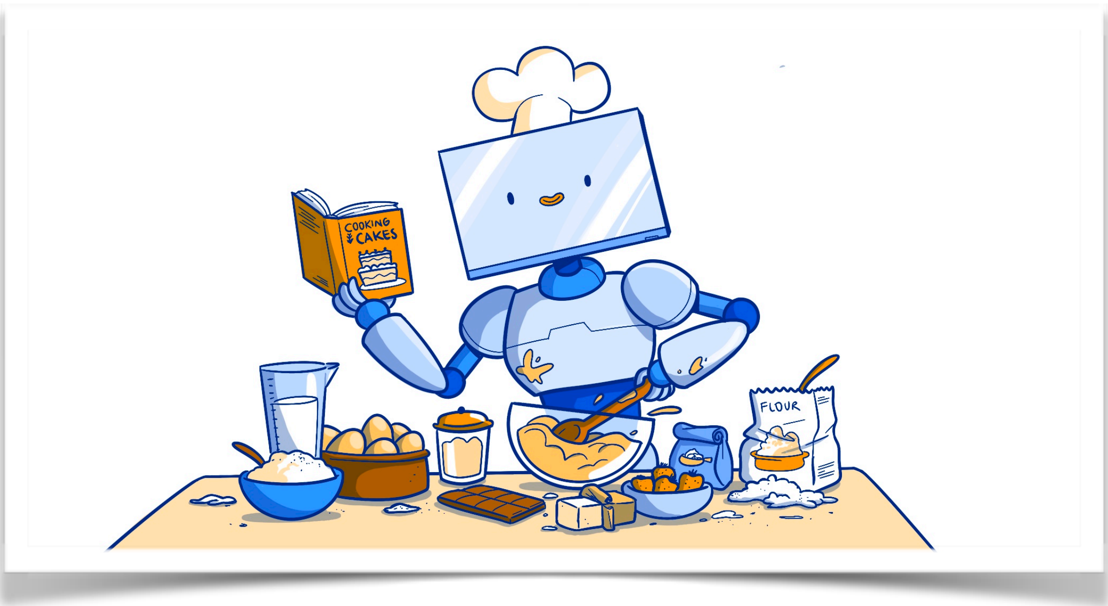
AI is a set of instructions which are written in a computer program (software). The instructions run a computer which performs mathematical tests on data. Which mathematical tests are used and what the result looks like depends on what problem we wanted the AI to solve. Humans then review the results the AI produces.



There are lots of different types of AI. The instructions that allow the AI to work are called an 'algorithm'.

What is an algorithm?

An algorithm is a set of clear instructions used to solve a problem. A recipe we follow to bake a cake solves the problem of 'I don't have a cake'. If we take the ingredients and follow the recipe, we should get a cake. An algorithm is a recipe, a set of rules, that only involve maths. By following this maths recipe, the algorithm solves problems.



IS AI THE SAME AS A HUMAN BRAIN?

Some people are immediately worried or concerned when we talk about AI. They imagine human like robots with super intelligence. These super AI robots do not exist and are not going to exist any time soon.

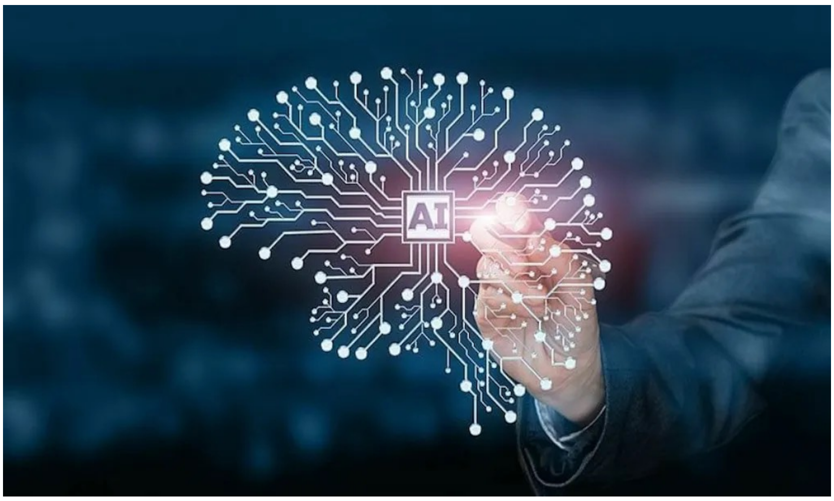
One of the most amazing things about humans is that we are so good at so many different things, like language, mathematics or using common sense when things don't go to plan. The list of our human abilities is enormous. There is not a single AI system that can do all of the things that a human can do. Most scientists think we are many away from having an AI that can match humans across all the skills we have.

However, there are AI systems that are 'better' than humans at some specific tasks. These are narrow highly focused tasks. These AI systems have what is called narrow artificial intelligence. Sometimes they are better than humans just because they can work faster. Sometimes they are better because they can see patterns in data that we can't because the data is too large or complicated.

"The list of our human abilities is enormous. There is not a single AI system that can do all of the things that a human can do."

But these narrow AI systems are only good at one thing. A narrow AI that can spot a cancer on an x-ray would not be able to tell if a blood result was normal.

In the world around us today we only have narrow artificial intelligence and this is the type of artificial intelligence that is starting to be used in healthcare today. This is the type of AI we will be taking about in the rest of this chapter.



WHO MAKES THE ALGORITHM (RECIPE) THAT THE COMPUTER FOLLOWS?

A computer follows a program that allows it to solve problems using mathematics. That program contains an algorithm, and it is this algorithm that gives the computer 'artificial intelligence'.

If we want to use artificial intelligence we need an algorithm, the recipe that the computer will follow. Who makes the algorithm?

Algorithms made by humans

The algorithm can be made entirely by humans. The computer is given a fixed set of rules that it follows. For example, if we make some software that contains the English dictionary and we put that software into a computer, we can then search on the computer for the definition of words. The computer is just giving us the definition we gave it but it is doing it much more quickly than we can by looking it up in a book. In the same way we can give the computer all the 'rules' for treating breast cancer. We write software with all the information on what treatments are used for different sizes and types of breast cancer. We can then give the AI details of a patient's cancer and ask what treatments the guidelines recommend. The AI would not be 'deciding' what treatment to give. It would be very quickly checking the patient's clinical data against the guidelines and rules that humans had agreed as evidenced by research.

Medical science has become so advanced that it is harder and harder for doctors to keep up to date with every new treatment so an AI like this can be helpful. These are rule-based algorithms, sometimes called expert systems. They have been around for decades. This is 'old fashioned' AI. The computer algorithms must be programmed into the computer by humans. Rule-based algorithms can be very useful for making simple tasks happen automatically. But they are limited in the types of problems they can solve because humans have to make the recipe (algorithm) for the computer.

Basic algorithm example

At their most simple rule-based AI involves the computer following basic instructions. These instructions can be described as IF and THEN.

For example, a simple AI which had the task of highlighting abnormal blood test results might follow rules like this:

IF result bigger than 10 THEN highlight in red.

That way humans can just check any numbers that are highlighted in red and might only skim the other results.

The AI is not creating new suggestions it is only doing a job we told it to do.

A rule-based AI like the one we talked about in the text that could help doctors decide what treatments to give patients with breast cancer would be much more complicated. However, at its heart it would still be following lots of instructions. For example, IF the age is greater than 75 THEN possible treatments are A, B, or C. But the complexity would be much greater as there would be lots of IF-THEN outputs which the AI would need to combine.

Algorithms made by computers – machine learning

Humans can write a computer program that allows the computer to learn for itself the best way to solve a problem. The computer is making the algorithm that it then follows. The computer makes the algorithm in a process called machine learning.

The computer is given lots of data (humans have to be very careful to only use good quality fair data). The computer looks for patterns in the data. Gradually it turns these patterns into mathematical formulas that allow it to solve the problem we set. It may find ways to group pieces of data together or to make predictions from data. Humans provide feedback on the answers. The AI will adjust the mathematical formulas to get better at solving the problem.

Video

This is an animated video on What is an Algorithm? To watch, open the camera on your phone or tablet and aim it at the QR code on the right. Centre the QR code and steady your hand for a couple of seconds. A link will appear on your screen, tap the link to take you to the video on YouTube https://www.youtube.com/watch?v=4dKoX7hBXlw.

How does this work in real life?

How could this work in healthcare? Let's imagine we are trying to find new ways to detect cancer at its early stages from routine blood tests.

If we made a rule-based algorithm we would need to tell the computer what was normal and abnormal, which results might be linked to cancer. This algorithm might help to review thousands of blood test results more quickly than humans could so it might be helpful. But it would be spotting things that we could spot if we only had the time to look carefully. It would not discover new clues in the data.

If we used machine learning, we would give the computer blood test results, including people whom we knew had cancer and people who didn't. The machine learning algorithm could be asked to predict which people in the data had cancer and which didn't. Humans would check the result and give feedback where the AI was wrong. The AI would then adjust the algorithm to try and be more accurate and humans would check it again.

This would be repeated until the AI was really good at predicting which people had cancer from looking at their blood test results. We would then give the computer data on a new set of patients and see how it performed. We would hope that in the end, the AI could flag up blood test results which showed warning signs of cancer. But the AI would have done this in a different way than humans would have thought of. It would almost certainly see patterns in the data that we humans could not see and use these patterns to make predictions.

Video

This is an animated video on What is Artificial Intelligence? To watch, open the camera on your phone or tablet and aim it at the QR code on the right. Centre the QR code and steady your hand for a couple of seconds. A link will appear on your screen, tap the link to take you to the video on YouTube https://www.youtube.com/watch?v=N4NZoTW8ekY.

Are computers really learning?

When we start talking about computers ‘learning’, is when some people start to get worried; scary science fiction movies spring to mind.

Machines or computers don’t learn in the same way that humans learn. What machines or computers are doing is using trial and error to get better and improve their performance. Because they can do thousands of mathematical calculations per minute, they are very quick at improving. This is what is meant by ‘learning’, the computers don’t understand the data. They are finding the quickest or most correct mathematical tests to use on the data to get the results the human has asked for.

With rule-based systems, when humans make the recipe (the algorithm) it stays the same unless humans change it. With machine learning the computer changes the algorithms and tests it thousands of times. The computer is trying to find the best algorithm to give the correct answers on the data it has been given.

Some AI looks like it has human intelligence

There have been recent advances in a type of AI called generative AI. Chat GPT and Google Bard are examples of generative AI. This type of AI uses machine learning. It is still narrow AI but it uses lots of algorithms added together to perform tasks. This type of AI is used to create (generate) things for example text, pictures or audio. It is very powerful, but it is still not thinking like humans. It is using probability to give the most likely answer to a question. It has been trained on over 300 billion words! However, that does not mean that it is definitely free from bias and it can give misleading results to some tasks.

This type of AI does not solve problems that relate directly to treating patients. It is not trained to look at results or scans or to make diagnoses. If it is improved further, it might be able to help with some administration tasks. For example writing letters between doctors and patients after a clinic appointment or making a summary letter at the end of a patient’s treatment. But at the moment it can still make mistakes and it is not approved for this type of work.

Real examples of rules based and machine learning narrow AI

Go is an ancient Chinese board game. The number of possible moves is even higher than the number of moves in a game of chess. It has been calculated that there are more possible moves in a game of Go than the number of atoms in our universe!

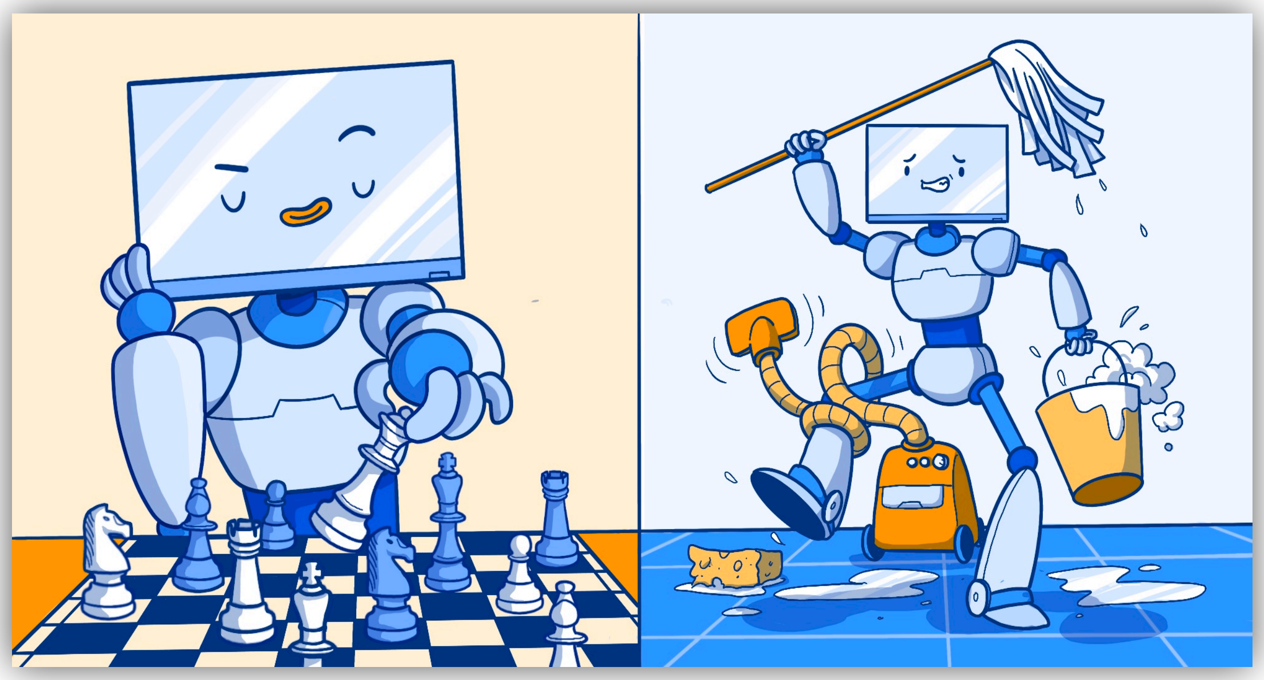
A company called Deep Mind developed a machine learning AI called AlphaGo which beat the human Go World Champion, Lee Sedol. The scientists that programmed the computer did not teach the computer how to play, they just gave it the rules of the game. The algorithm then evolved by playing thousands of matches against amateurs and then professional players. The computer analysed what combinations of moves led to a win and which lost. It gradually improved until the AI was good enough to beat the world champion.

Narrow AI can only solve a narrow set of problems.

The example of an AI algorithm being better than humans at a complex game like Go can raise concerns about the power of some of these AI systems. But remember, these narrow AIs are only good at the one thing.

If we gave the AlphaGo AI an IQ test it would fail! It would not be able to answer any of the questions and would have an IQ of zero!

We could build an AI to answer the IQ test and over time we could probably get it to pass the test with the highest possible mark – an IQ of 201. This could look impressive, but the AI would not really have a high IQ. If we tried to use this AI for anything else, it would be useless.



	Rule-based AI	Machine learning AI
Who makes the rules for the algorithm receipt?	Humans	Computers
Can we understand what the AI is doing?	Yes	Not always
Can the AI make new discoveries that humans hadn't realised?	No	Yes
How safe is the AI?	If the humans have been careful making the algorithm in the AI, then the risks are low.	The safety of the machine learning AI depends on how good the data was that was used to make the machine learning AI. If there were problems with the data, then the AI would not be safe.
Examples	An alert system that indicates abnormal test results.	A computer system that finds patients who might have a heart attack in the next year by looking at the data from a GP practice.
	An alert system to warn doctors when they are prescribing the wrong medication.	A computer that can find when cancer is present on a scan.

CHAPTER 7: PROBLEMS WITH USING AI IN HEALTHCARE

Now we understand what AI is and how it is created, in this chapter we will learn about the problems that can happen when we use AI in healthcare.

Most of us accept and use AI everyday, although we may not realise it. Having AI select which pages to show you from an internet search is something many of us don't think about. There are risks with this trust, how do we know we are being shown the 'best' or 'most accurate' internet pages by the AI? But most of us are not too concerned about this. However, when people think about AI in healthcare then they may be more worried.

As we have learnt in chapter 6 machine learning AI has the possibility to be more powerful than rule-based AI. But if we use an AI that is based on rules that humans make, then it is easier to check and understand if it is safe. Being sure machine learning AI is safe is much more difficult. In this chapter we will look at three problems that we need to overcome if we are going to use machine learning AI safely in healthcare: generalisability, explainability and data quality.

AI TO IMPROVE EYE CARE

Google have created a machine learning AI to look at photographs of the back of the eye (the retina). The AI can spot diseases from looking at these photos of the retina. For example, diabetes is a disease that can damage to the back of the eye. From the retinol photos the AI can identify which patients are developing eye damage due to their diabetes. As there are millions of people with diabetes in the world the potential for this AI to help is enormous. The algorithm took less training than a doctor (in the UK a consultant ophthalmologist will have had 5 years at medical school and then 8-10 years of further training until they become a consultant). It also takes less time to look at the photos and doesn't need to stop to eat or sleep unlike human doctors!

But there can problems.

There have been problems with this AI when it has been used in different countries. Eye clinics are run differently in different parts of the world and the retina photo AI didn't work as well in all places as the way the retinal photos were taken varied as did the types of eye diseases seen in different clinics. AI in healthcare needs a lot of work and testing so it can be used safely everywhere.

Scientists can often get good results from an AI in a laboratory or one hospital. But the results are not as good when the AI is used with real-world data from lots of different places. This is the problem of generalisability.

GENERALISABILITY

Generalisability is how well an AI works when you use it in different places from where it was made. When you use the AI somewhere else it will be using data that is slightly different from the data that was used to make it. The data might be different because it involves different types of people. For example if the algorithm was made using data on adult patients it would not work if you tried to use it in the children's hospital. But also, even if you use the algorithm on the same type of people how the data was collected could be different. If we have an AI that works on x-rays in one hospital it might not work as well in another hospital just down the road if they use different x-ray machines because the x-ray images might not be the same.

Covid-19 and AI

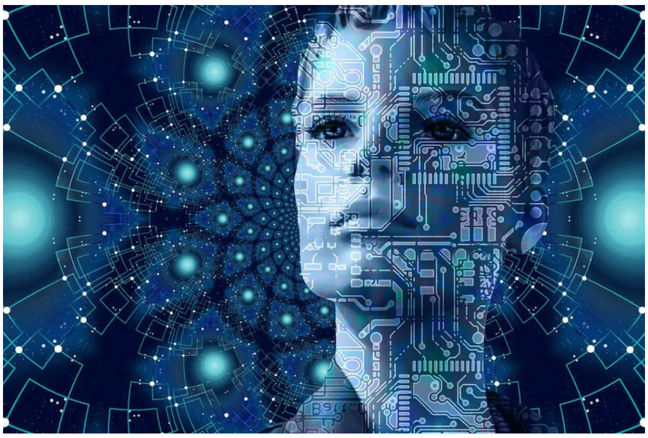
At the start of the pandemic hospitals were overloaded with patients and sometimes it took several hours to get a rest result for Covid-19. Many patients had chest x-rays as they were very unwell so researchers tried to create AI which could diagnose Covid-19 on chest x-rays.

Researchers could create AI algorithms that worked well on small sets of data from one hospital. But there were very few of these AI algorithms that worked well in lots of situations. Researchers have reviewed many of these algorithms and one of the biggest problems seems to have been that the data that was used to train the AIs was biased. This meant that that the algorithms did not work well in different hospitals. They were not generalisable.

Explainability

Another problem with machine learning is that we cannot always explain the results.

If the computer has made the algorithm, how do humans understand what the AI has done? This can be difficult, and it is why machine learning AI has been called a "black box". We know what data goes into the AI and we see the answer come out of the AI, but we can't always know what went on inside the AI. Black box is a way of saying there's no window that allows us to see exactly what mathematics the AI is using. This means that if the AI comes up with a wrong answer, we might not know why. We might not understand its recipe. Scientists are working on ways to help us understand what the machine learning AI has done but these are not yet perfect.



Data quality - Bad data makes bad AI

Let us revisit what we learnt in the sections on data and think about the problems we can have with data. The data we use to make the AI will have a big impact on how well the AI works in real life. Bad data will make bad AI. Bad AI can mean it doesn't work very well. It might give us the wrong answers, or only works well for some people but not others, so it is not safe for everyone. Here are some examples of how problems with data caused problems making good AI.

PREDICTING WHICH PATIENTS NEED TO STAY IN HOSPITAL

A machine learning algorithm was made to predict which patients with serious chest infections (pneumonia) could be sent home safely and which needed to be admitted. The data showed that people with pneumonia who also had asthma were less likely to die from the pneumonia. This was spotted by the computer. It used this pattern in the data in its algorithm. That meant that AI was more likely to suggest sending people home who had asthma.

But there was a problem with the data. It did not show why patients with asthma were less likely to die. The data did not show that because doctors were very worried about patients with asthma who had a serious chest infection, they gave them even more intense treatment. It was the intense treatment that meant patients with asthma were less likely to die, it wasn't the asthma itself that was protecting them. Sending patients with asthma home would have been a big mistake.

What is interesting here, and happens a lot when working with AI, is that the AI teaches us something. In this case if we treated some patients with serious chest infections who don't have asthma with more intensive treatment, could we get better outcomes for more patients?

IS THIS A MOLE OR IS IT SKIN CANCER?

AI systems have been made to try and diagnose a type of skin cancer called melanoma from photographs. Some of these AI appear to work well. But do they work as well for everyone?

The data used to make some of these AI algorithms does not include all people in a way that means the AI works equally well for people of all skin colours. Melanoma is more common in people with white skin so there are more photos of melanoma in white skin than in black skin or other skin colours. This means when AI was being made to look for melanoma, it will have more data about melanoma in white skin. When the AI is then used in the real world there is a risk that it won't work as well for all skin colours. The problem here is not the AI. It is that the data used to make the AI did not have enough data on people with different coloured skin.

Researchers working to develop AI to diagnose melanoma are working to make sure there is enough data for people with all skin tones.

Spotting breast cancer on x-rays

There are lots of people trying to create AI systems that can help diagnose breast cancer from x-rays of the breast (mammograms). But human breasts are not all the same. Some have more fat in them, others have more fibrous tissue in them. Women from all backgrounds can have different types of breast tissue but fibrous breast tissue is more common in women of Chinese ethnicity.

This means that if you are making an AI to find breast cancer on x-rays you need to consider if you have enough data on all types of patients and where you will use the AI. If you made the AI only using data from the NHS there is a chance you will not have enough examples of fibrous breast tissue so the AI would not work so well for women with fibrous breasts. This would mean that it might not work as well for women of Chinese ethnicity. The same problem could happen if an AI was made only using x-ray data from China. The AI might work well for most women of Chinese ethnicity, but it might not work as well in breasts which had more fat in them, such as in women of white ethnicity in the UK.

What can we do to make sure AI is safe?

There are many tests being used to check if the AI 'does what it says on the tin' but there is not a single test or measurement that tells us if an AI is working well. We need to use a variety of different tests and measurements to decide if the AI does what we want and does it safely in many different settings.

When you buy new tyres for a car, they have a label on them. The label gives you the results of different tests. How the tyres affect how much fuel your car uses, how noisy the tyres are on the road and how good they are when you need to brake in wet weather. You decide what you think about these ratings which are measuring safety and you add in what the tyre costs so you can decide if it is value for money. It is very similar with an AI product. We need several measurements to try and have an idea as to whether the AI works well, if it is safe and if we would want to use it.

How is AI regulated and approved?

At the moment, AI is being regulated and approved in the same way medical equipment is approved. That is OK for rule-based AI. Rule-based AI doesn't change unless humans deliberately change it, and they would know and understand those changes.

But this doesn't work in the same way for machine learning AI. As the machine learning AI is shown new data it will change its algorithm. Usually once a machine learning algorithm is being used in the real world it must be retrained by humans who will give it new sets of data based on where the algorithm is being used. This means the algorithm will change, maybe only in very small ways, but it can be hard for humans to understand how the algorithm has changed or why. There are also machine learning algorithms that retrain themselves automatically as they 'work' rather than needing humans to give them sets of training data. These types of AI are not yet common, but we need to think how we would ensure they were safe.

This all gives us a problem when we try to test and regulate things that are changing. Different rules and regulations are needed than the ones we have for normal medical devices. In the UK, the British Standards Institute (BSI) is developing new standards for AI in healthcare and the National Institute for Health and Care Excellence (NICE) has already published standards for what evidence is needed to use digital technology in healthcare.

We want to make sure AI is safe but delivering healthcare is becoming harder and AI could help us (we will see examples how AI might improve healthcare in the next chapter). We need to find a balance between allowing AI to develop quickly so it can help us make healthcare better but balance that with safety and being sure no harm is caused.

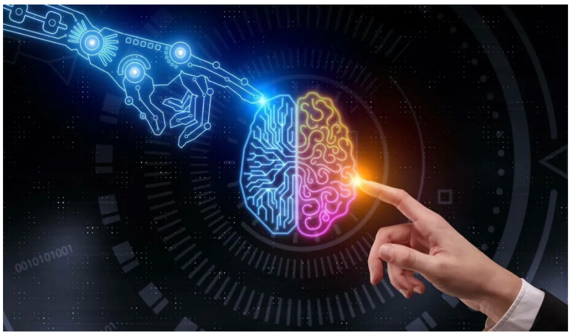
CHAPTER 8: THE BENEFITS OF USING AI IN HEALTHCARE

In the last chapter we started to learn about some of the problems that we have to overcome if we are going to use AI in healthcare. Some of these problems are difficult to solve. You might ask, 'why bother?' 'Why do we need AI in healthcare?' In this chapter we will look at some possible benefits and then you can decide for yourself whether you think AI in healthcare is a good thing or not.

IT'S NOT ALL BAD NEWS – AI CAN CORRECT HUMAN MISTAKES

Knee pain is common. One cause of knee pain is arthritis of the knee. This is where the lining of the knee joint wears out and becomes painful.

Surgeons use x-rays to help diagnose arthritis and decide if a patient would be helped by knee replacement surgery. Diagnosing arthritis on x-rays is a skill that doctors develop during their training. The arthritis often develops in a standard pattern that is easy for experienced doctors to find.



Not everyone who has pain in the knee has arthritis. Sometimes no cause for the pain is found and people must live with long term knee pain.

Researchers noticed that more patients who were black were diagnosed with chronic knee pain. These patients had x-rays which the surgeons did not think showed arthritis. They also noticed that fewer people who were black had knee replacements, so they tried to find out why.

The data about these patients was reviewed using AI. The scientists found that the surgeons were right about some of these patients and wrong about others. Some of these patients did have arthritis when the surgeons thought they didn't BUT it was not the normal pattern of arthritis that surgeons are used to diagnosing. Because surgeons had not realised you could get this different pattern of arthritis, they didn't know to look for it ,leading some patients having their diagnosis of arthritis missed and so they didn't get a knee replacement.

Humans had been making this mistake for years. We needed something to look at the data from a new perspective. The AI wasn't studying the data in the way humans would and this meant it could find something we had been missing.

This shows us that whilst there can be problems with AI there are also problems with humans. We make mistakes and we don't always realise why we are making them. AI gives us a chance to notice things in the world around us that we haven't yet seen.

There is a risk of bias with AI if the data is biased but there is also a hope that AI can challenge some of our bias and make healthcare fairer for all patients.

Podcast

This podcast is a conversation between Jonathan Gregory, a doctor, and Danny Ruta, an AI expert, about using AI in healthcare. To listen, open the camera on your phone or tablet and aim it at the QR code on the right. Centre the QR code and steady your hand for a couple of seconds. A link will appear on your screen, tap the link to take you to the podcast on YouTube https://www.youtube.com/watch?v=tPTSPphQRNA.

Examples of how AI might improve healthcare

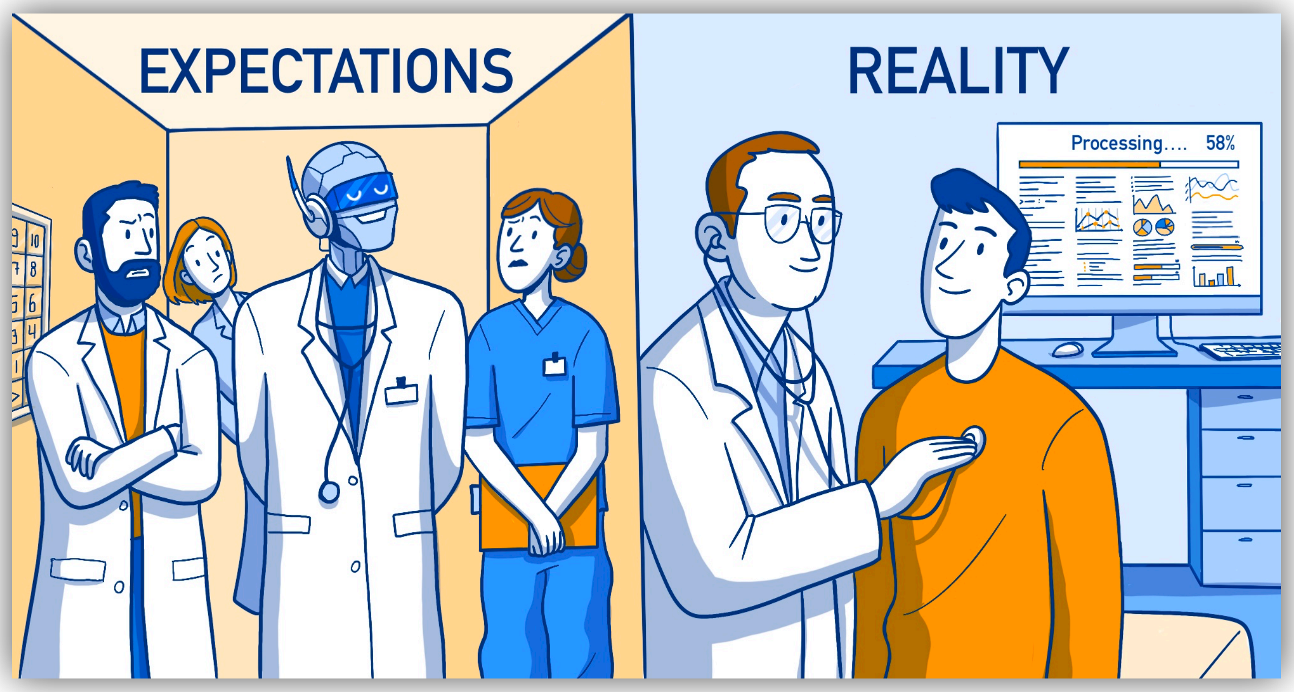
Problem	How might AI help healthcare teams?
There are not enough x-ray specialist doctors, so patients wait a long time for scan results.	AI might be able to check all the x-rays and scans and find those that are normal. It could create an automatic result that was sent to the patient and their doctor the same day as the test was done. Doctors could focus on abnormal x-rays, or scans that the AI was uncertain about.
Doctors sometimes miss an abnormality on a scan or blood test.	AI might be able to double check the doctors. By reviewing all the tests, it could alert humans to double check if the AI thinks something is abnormal, but the human doctor hadn't noticed.
Doctors spend a lot of time checking test results which are normal.	An AI could automatically tell patients they had a normal test result and file the result in the electronic medical records or book an appointment for clinic.
Nurses and healthcare assistants spend a lot of time checking hospital patients blood pressure and temperature.	AI can assess the results from devices that automatically measure temperature and blood pressure. The AI could alert the nurses to problems but if everything is OK the nurses can carry on with all the other work they have to do.
Booking appointments for clinics and scans takes a lot of time and effort for human staff.	AI will be able to automatically schedule clinics and scan appointments and try and help them be booked so they run on time by not being over or under booked.

Will AI replace doctors and nurses?

Currently all the plans for using AI in healthcare are for it to help all the people who help us with our health. We are not about to have AI replace doctors, nurses, midwives, physiotherapists, occupational therapists and all the other people who help us to get well and stay well. The hope is that AI will soon be helping them with their jobs. Helping them to provide higher quality care and to manage staff shortages by AI helping to prioritise what is important.

When AI is going to be used to help healthcare staff decide which treatment a patient should receive, a human will have the final decision on what to do.

Over time AI will improve, we will start to trust it and understand when to use it and when not to use it. When automated lifts became more common in the 1920's many people did not trust them. It took years before people felt safe getting into a lift without a lift attendant. Today most people get in a lift without thinking or worrying about how it works and if it is safe. It is quite likely that over many years patients and their health professionals will trust AI to do more for us.



How AI might free humans to be more human?

Computers have made a big difference to healthcare. Many of these changes have been for the better but computers can cause problems.

Before computers were commonly used, an appointment with a doctor would involve the doctor asking questions and listening to the answers. The doctor would also be trying to build a relationship with the patient. They would also look at the patient's body language to help understand how a patient was feeling. One of the reasons people complained about doctors' handwriting is that they would write very short notes very quickly at the end of the appointment so they could listen carefully to the patient when they were speaking.

Once computers came along the doctors started to spend more time looking at the computer and typing things into it than looking at the patients. This has meant that some patients do not feel like the doctor has listened to them properly. It has also caused some doctors to not enjoy their work as they want to be 'with' their patients, not looking at a computer.

AI might be able to solve this. It might help us to bring the benefits that computers offer but reduce the negative impact they can have on doctor and patient communication.

As long ago as 2006, software was used which could take what the doctor was saying during a conversation with a patient and start to fill in the paperwork. For example, if the doctor was going to refer the patient to check if they had cancer, the computer would start to complete the referral letter from what the doctor had said. This reduced the paperwork doctors had to fill in after each appointment.

Unfortunately, the methods used in 2006 could not work in every situation so it did not become widespread in the NHS. Now we have a better understanding of AI, researchers are again trying to solve this problem. Some very large companies are trying to make AI which will mean that doctors won't have to type into the keyboard of a computer. They are trying to make an AI system that will pick words out of the conversation and help the consultation by bringing up test results automatically so the doctor can discuss them. It will start to automatically order tests the doctor thinks are needed and write the letter that is sent to the patient and their GP after every appointment.

Computers were brought into the NHS to solve some problems but they created others. AI might help us to have the benefits of computers but stop them getting in the way of patients and doctors having proper conversations

OVER TO YOU

The use of new technology in healthcare is always difficult and problems will occur. Whenever there has been a brand-new operation, medical device, or drug treatment there have often been problems to overcome. In just the same way, the use of AI in healthcare will cause some problems.

We think that none of the possible problems on their own are a reason not to use AI in healthcare. Being a doctor, nurse, midwife, pharmacist or physiotherapist is harder now than at any other time due to the huge advances in medical science. There is now too much for any human to know or to keep up to date with. AI can help people do this. It can release professionals from administration tasks giving them more time to spend with patients and supporting people to stay healthy or to get better. AI can study data in new ways and help humans to better understand health and disease.

We think that everyone needs to be careful with AI. We need to work hard to make it safe by understanding how it might go wrong and how to prevent this from happening. AI offers a real opportunity to deliver better care for more people around the world. A careful approach will allow us to receive help from what AI has to offer while minimising the risk of harm.

What do you think? Having thought about data and AI do you agree or are you worried? Will you put your data into a healthcare app to help you manage your health? Would you let doctors use AI to help them to help you with your health and well-being? Will you agree to share your health data so that research has results that work for everyone without bias?

CHAPTER 9: WANT TO KNOW MORE?

If you'd like to know more

We hope you've enjoyed reading this guide. You might want to know more so here we have a list of resources that the team have used and found helpful. We have tried to make sure that all the links work and that the material is correct, but things change over time so if any links are broken or the resources don't age well, we are sorry and please let us know.

If you find any other resources that you think are helpful, please get in touch and let us know and we can add them to the list.

Books

These books have had a big influence on this guide.

- The book by David Spiegelhalter is fantastic. We have rewritten some of the examples he gives to make them quicker and easier to read. His book covers the issues in more depth and with many more examples.

  The Art of Statistics Learning from Data, David Spiegelhalter ISBN 978-0-241-25876-7

- The book by Eric Topol helps to relate how data and AI might be used to change the way we receive medical care in the future.

  Deep Medicine, Eric Topol ISBN 978-1-5416-4463-2 (hardcopy) and 978-1-5416-4464-9 (ebook)

Available Online (if you are reading this on online then click on the link)

- Dstl Biscuit Book and Dstl Crumbs! Understanding data. Both available for free from this websitehttps://www.gov.uk/government/collections/dstl-biscuit-books
- This guide has more examples of data in healthcare and helped us produce our guide. Data science a guide for society. Sense about science
- Living With Data is a research project which has worked with the public to explore what it's like to live with data, AI and automation https://livingwithdata.org

More in depth

We think the resources below cover things in more depth but because of that they may take more time to work through.

Books

- AI in Healthcare: Theory to Application, Dr Sandeep Reddy
- Machine Learning for Absolute Beginners, Oliver Theobald

Online

- AI for Healthcare: Equipping the Workforce for Digital Transformation. This is a free online course developed by the University of Manchester and Health education England. It takes about 10 hours to complete and takes you deeper into Ai in healthcare.
- The Health Foundation. How do we get the best out of automation and AI in healthcare?
- The Ada Lovelace institute report into data driven technology and health inequality

  These are both reports from well-known and trusted organisations which explore how data and AI may impact health.
- Statology is an online statistics and AI courses and tutorials. These courses are not free but some of the content can be read without payment. https://www.statology.org/about/

ACKNOWLEDGEMENTS

We would like to thank everyone who has contributed to creating this guide.

PROJECT TEAM PUBLIC CONTRIBUTORS

Amrapali Amrapali
Vernon Bailey
Christine Chaloner-Sandell
Saran Green
Tene Green
Pamela Gregory
Suzanne Kelly
Mark Minott
Steven Murphy
Michael Njegomir
Victoria Nnatuanya
Othman Mohammad Noor
Reshma Punjabi
Emma Robertson
Serdar Savasir
John Schofield
Liz Schofield
Suzanne Steer
Shanice Thomas
Anthony Austell Turner-Truro
Sandra Tymms
Pournamy Vema

PROJECT TEAM EXPERT CONTRIBUTORS

Kelly Gleason
Jonathan Gregory
Kerlann Le Calvez
Danny Ruta
Martha Martin
Kelly Mousa
Ronnie Davies
Matt Williams

People who have reviewed and commented on the guide or have had conversations and discussions with us that have shaped the guide.

Hatim Abdulhussain
Phil Booth
Dorota Chapko
Tom Clutton-Brock
Alan Davies
Ang Davies
Ilianda Eleftheriou
Glen Hart
Francis Hooley
Charlie Huins
Halle Johnson
Chris Kelly
Julia Letts
Rachita Mallya
Kabelo Murray
Hina Pandya
Maria Piggin
Stefanie Posavec
David Spiegelhalter
Piotr Teodorowski
Vishaal Virani
Kieran Zucker

OUR FUNDERS

Imperial Societal Engagement Participatory Research Seed Fund Award
The Imperial Experimental Cancer Medicine Centre
The Cancer Research UK Convergence Science Centre
The NIHR Applied Research Collaboration for Northwest London
The NIHR Imperial Biomedical Research Centre – Surgery & Cancer Theme
The NIHR Imperial Biomedical Research Centre - Patient Experience Research Centre

OUR COLLABORATORS

Cognitant ~ illustrations and animated videos
Evolene ~ cover design
Healthwave ~ UK wide public consultation of content

OUR JOURNEY INTO AI

This mandala is a representation of our learning through co-creation of this guide. We learned a great deal about data and AI but the experience taught us much more than that. We accepted that healthcare is changing and that it is important that we are part of the conversation so that we can make informed decisions about our health and the health of the people we love. We hope the guide has helped you to better understand data and AI so that you can feel more confident when communicating with healthcare teams and engaging with the changes that data and AI are bringing to healthcare.