ARTIFICIAL INTELLIGENCE & DATA SCIENCE IN HEALTH

Educational resources made for the public by the public

INTRODUCTION

Imagine…you have diabetes and your doctor suggests an app that will use AI to help manage your disease. What questions would you ask to decide if you agree to trust this app with your data? What might you want to know about the AI? Who would you ask? Could the GP tell you what you wanted to know?

We have co-produced an entry level educational resource around data science and AI. Content has been co-produced by members of the public with the support of experts. The resource has been made by the public for the public.

Podcast

To listen to a conversation between Stephanie Posavec, an artist that works with data visualisation, and Reshma Punjabi, a public contributor to this resource, open the camera on your phone and aim it to the QR code. A link will appear on your screen follow that link to the podcast on YouTube.

SCAN ME

CONTENTS

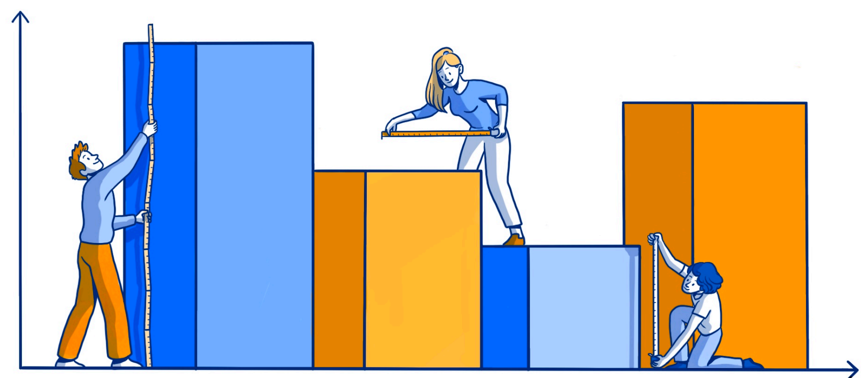
CHAPTER 1: WHAT IS DATA?

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

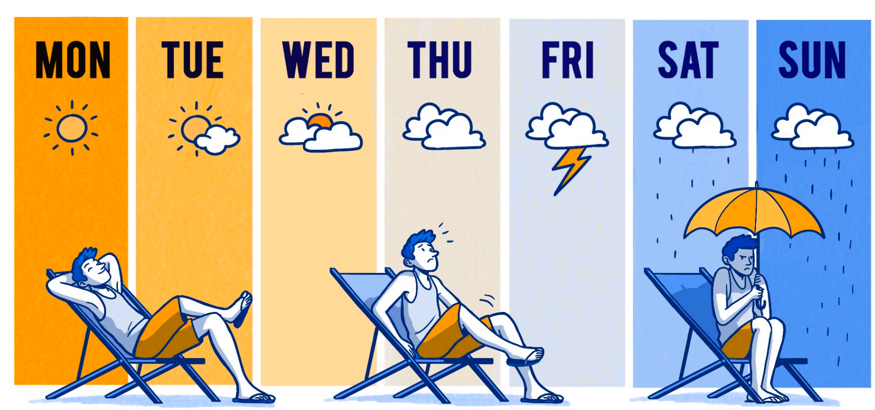
There are lots of different ways of describing data. We will think about some of the common ways so we can all have a good understanding of all the different types of data.

As we have already mentioned, there are lots of different types of data. For example, words, numbers, dates, and times. There is a lot of 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.

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! 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 doctors 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 really hard 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 2cm x 5 cm' , Blood pressure 120/80.

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

There would also be *data in your medical records* for example 'is not allergic to any medicines'.

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 2cmx5cm on the front of the left elbow. Looks like eczema. Try low dose steroid cream'

The GP weigh's up this information to make a diagnosis.

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

Diagnosis = Eczema

Location = Left elbow

Treatment = steroid cream

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

'I saw the GP about the rash on my arm. She asked me some questions, had a look and sent me to get some cream from the chemist'.

Information which can be used to identify someone includes their names, date of birth, full postcode, address, email address or phone number. If the data does not have any of these details, then it is very hard to identify individual people.

There are always risks that people could illegally work on the data to identify people but without these details it is very difficult.

Can people be identified from data?

When people talk about data, words that are used include 'anonymised' and 'pseudo-anonymised' data. What do these mean and how could it affect you?

Anonymised data is where information about a person's identity has not been included. The data does not include names, date of birth, full postcode, address, email address or phone number.

Anonymised data may still capture some details about the person but these are not enough to identify them. For example they may collect data on age group "people between 60 and 65". Or they may collect what region of the country they are from "live in Birmingham'. But it is extremely difficult to identify people from this data.

Pseudonymised data is data where people have tried to protect the identity of people.

A lot of the data that could be used to easily identify someone is removed. For example date of birth is replaced by "age at the time of answers", only the start of the postcode is kept instead of the whole postcode. It is still hard to identify a person from pseudonymised data.

"Personal Identifiable data is data that could allow you to be identified from data. Examples of this type of data are your name, date of birth, address, postcode, email address and phone"

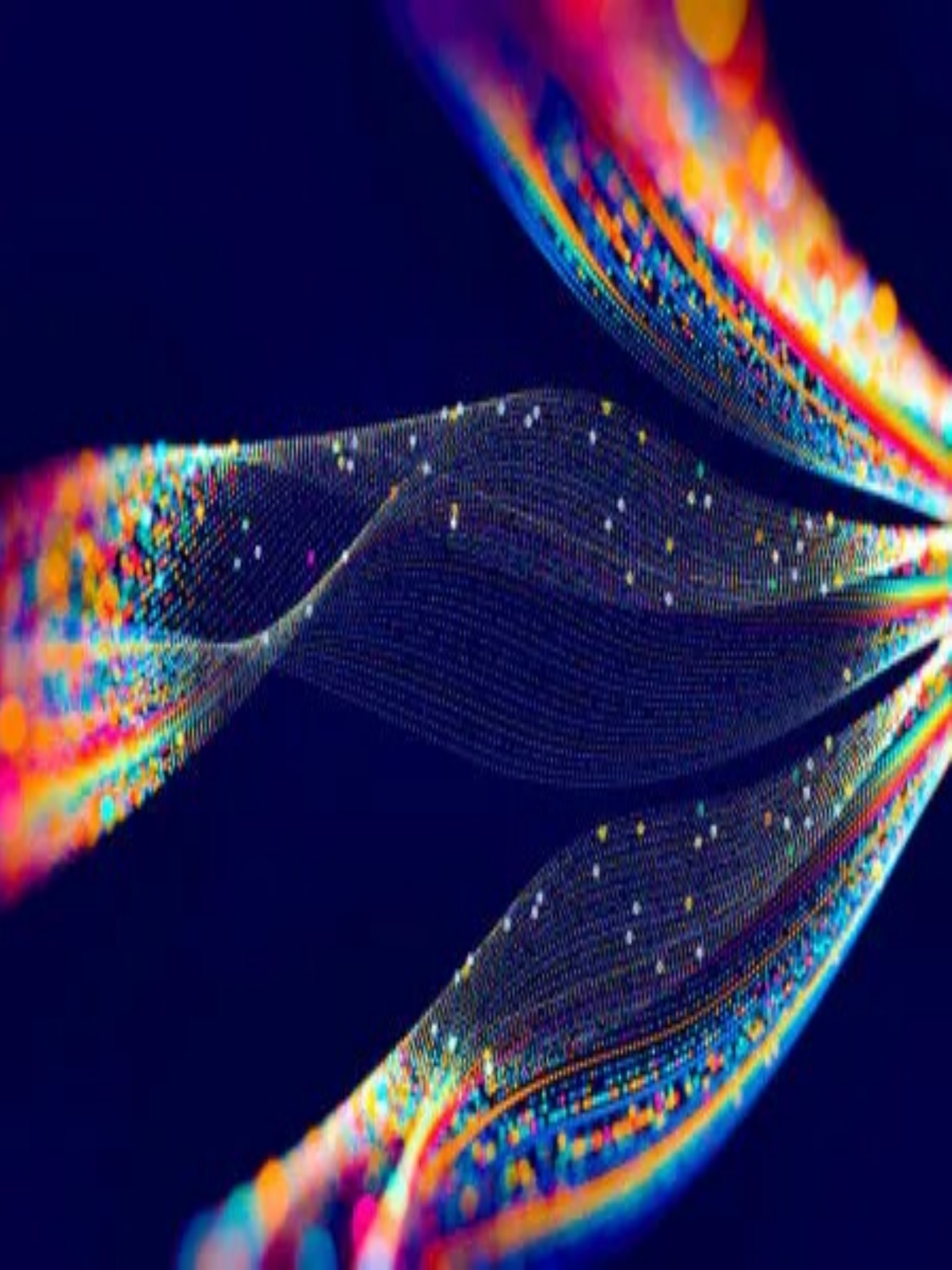
So why can’t all data be anonymised? Fully anonymising data is something that is not always easy to do. Data such as age and postcode can be used to check if the data represents the whole population (see later examples of where having data that doesn't include everyone caused problems).

For example, if we wanted to improve care for teenagers with diabetes, we need to know that the people in the data are teenagers, we need their age. Or maybe we want to improve GP services in a city, then we need to be able to check that the data we have covers the people who are living in that city.

Keeping those pieces of information helps us to do good things with data, but it also makes the people in the data vulnerable as someone could try and identify them from the data. This is illegal, and difficult, but that doesn’t stop some people from trying.

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

- Who is collecting my data?
- What do they want it for?
- How will my data be used? What does the small print say about how my data is going to be used?
- Who will be able to access my data? Where is my data being stored?
- Is my data secure?
- How long will my data be kept?
- What will happen to my data once it has been used?
- Is my data going to be used for other purposes other than the product or research I am signing up to?
- How personal is the data that I'm giving? Could I be identified from the data? Will the data be anonymised?
- Can I trust this person / company / organisation with my data?
- What is the benefit for me in giving my data? Are there benefits to anyone else (a company, the NHS, other people)
- Can my data be misused?
- Will this data be freely available to be used by others?
- Are there any standards or laws that the organisation/individual should be following? For example, GDPR.
- Where is the data stored? What is the environmental impact of this?



CHAPTER 2: TYPES OF DATA USED IN RESEARCH

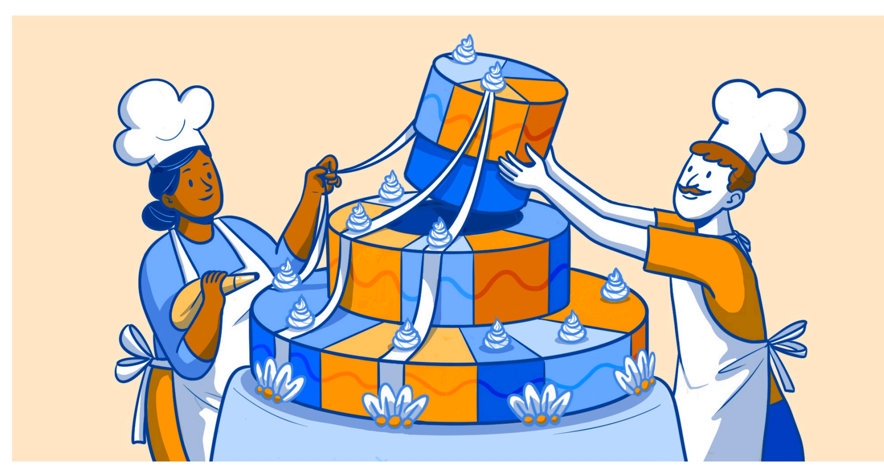
Building on Chapter 1 we explain how data is linked, and also some common terms you will hear about it.

Is the data joined up or "linked" to other data?

Joining different pieces of data to each other helps us to have a deeper understanding. This is really important in healthcare when we need to really and try and understand patients and their health. We can do this with Linked data.

Linked data is off interest in healthcare. Joining different pieces of data to each other helps us to have a deeper understanding. This is really important when we need to better understand patients, their health and their care needs. Linked data can help us do this.

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 patient in that area.



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, makes research cheaper and more related to the real world.

Open data 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.

Big data is a phrase you will hear in the media. What is it that makes some data big and other data small?

Big data is a way of describing data where there are so many millions of individual pieces of data that normal computing software or hardware cannot process it.

'Big Data' is actually a term from the past when only a small number of people had computers powerful enough to cope with data that was so large. Now computing power and the methods to work with data have improved, what was once thought to be big data is now ‘normal’ data. So, when we talk about “big data” these days, we really mean big…millions and even billions of pieces of data in one dataset.

Big Data studies are now usually about finding trends and patterns in places where nobody has even thought to look. .

“Big Data studies are now usually about finding trends and patterns. ”

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 all 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. But there is a risk. If you do not really understand the real-life data, you will create synthetic data which is not really like real life and so the results of anything you use the data for (for example research or to make a new piece of technology) will not work in the real world in the way you thought it would.

Fictitious examples of big and small data.

Your GP is doing a research project and they are looking at the result of one type of blood test result for 40 patients. Is this big data?

This is not big data, this is a small study. It may still be very important but the amount of data is not going to be very large and only simple data skills will be needed.

Your GP is doing a research project and they are looking at the result of one type of blood test for 1000 patients. Is this big data?

This is getting more difficult. 1000 results will be getting more difficult to manage, but it still isn't big data as there are not a lot of different categories (dimensions) in the data.

Your GP is doing a research project and they are looking at the result of 20 types of blood test results for 5000 patients over the last 10 years. Is this big data?

This is sounding like a big data project. Each patient will have a lot of results and it will require expertise to be able to look after this data and use it properly.

Remember that bigger is not always better.

The quality of the data and the skills of the people using the data are really important.

Well, what's the point of creating data that looks like the real world, haven't we got enough data already?

Well one reason is that sometimes we may have the data but don't want to release it to the analyst because it contains very personal information.

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) set of data, which could then be released for analysis as there are no details of any real people in it.

This can be helpful in terms of sharing data securely, making sure that people's privacy and confidentiality can be protected. One of the risks however, if you do not really understand the real-life data, you will make synthetic data which is not enough like the real data so any results from work using that synthetic data could be misleading or just plain wrong.

There are lots of ways to make synthetic data. Artificial intelligence (spoiler alert: more about that in Chapter 3) can make up data that humans can't tell from reality.

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. All these different types of data can be found in patient's medical records and can be used in medical data research. With so much data what types of mistakes can happen and how would we spot any problems? In this chapter we will look at some of the problems that can happen. We will also look at bias in data. What is bias in data and what problems can it cause?

Missing data

It is easiest to think about this with an example. Imagine we work for the NHS and we want to give money to GPs to make sure more patients who have high blood pressure have it checked regularly. Having a recent blood pressure result can let the GP know if the person's blood pressure is well controlled or not, and whether they need help to get it back to normal so they don't have health problems in the future.

We ask the GP Practices to tell us about their patients who have a diagnosis of high blood pressure recorded in their notes, so we can see if they have had their blood pressure checked recently.

GP Practice A sends us data for 100 patients who have a diagnosis of high blood pressure recorded in their notes. 90 of these patients have a blood pressure result but 10 of these patients don't have a recent blood pressure result in the records. They have "missing data" for the ten patients who have not had a recent blood pressure check.

GP practice B send us data on 80 patients, who have a diagnosis of high blood pressure recorded in their notes.

All of them have a recent blood pressure result recorded in the records. At first it seems like GP Practice B are better at checking the blood pressure for their patients who have a diagnosis of high blood pressure as they have results for 80 patients out of 80.

GP Practice B could have a different problem. If they had patients who had high blood pressure but the diagnosis had not been saved in the records, they would have missing data. Patients who should be in this list because they have high blood pressure would have been missed off the list because the diagnosis was not recorded in their records.

In the real-world GP practices are very good at recording patients diagnoses and checking blood pressure results of people with high blood pressure. But hopefully this example shows you that is can be hard to compare data from different places and you have to be careful as missing data can be hard to spot – because its 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, but it is 1.6cm because you didn't read the units carefully.

Usually this is easy to spot as no one is 1.6cm 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 160cm tall but you make a mistake, your finger slips on the keyboard, and you type 180cm tall. It is possible that someone is 180cm 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.

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 notes after the date recorded as the date of death.

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 or had low mood after losing this dog).

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 our 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 it can be biased. That can mean that 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 compared to another. Both of these mean that the data does not reflect how the world really is. We have to check 'is the data being fair to everyone'?

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.

People left out of the data

We have to make sure that we have data that represents everyone that might be affected by our results.

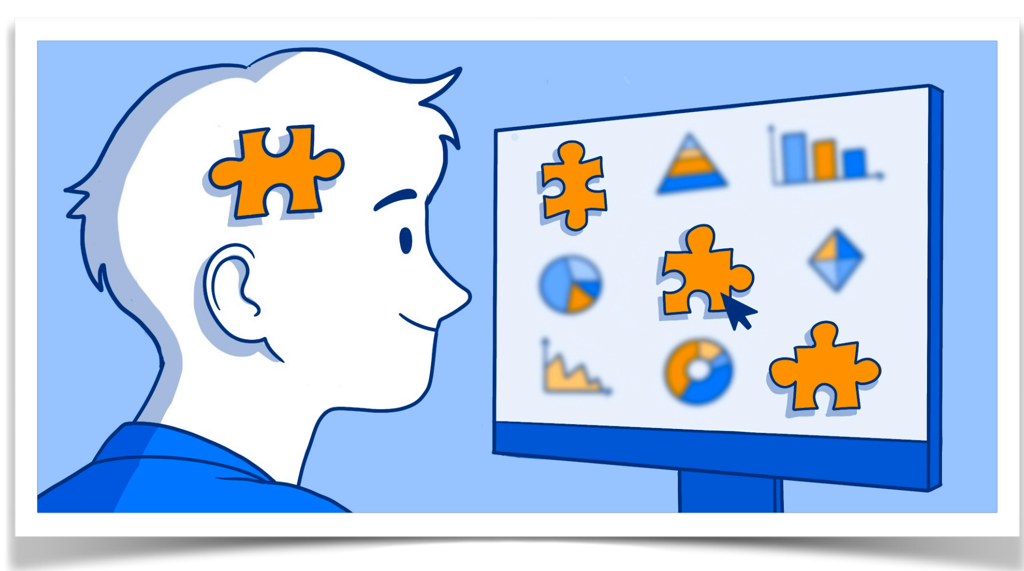
Let's say, in a large city in the USA, they wanted to be better at fixing pot holes in the roads. One problem was locating pot holes so they built an app that people could have on their smart phone. The app on the phone detected when people drove over bumps in the road and reported the location to the pot hole repair team. Owning a smart phone is more common in wealthier areas of the city, this meant that in poorer areas of the city fewer people used the app. This meant less data was recorded and these areas had fewer pot holes repaired. The people in poorer areas of the city were left out of the dat because of the way it was collected, using an app on a smart phone.

Types of bias

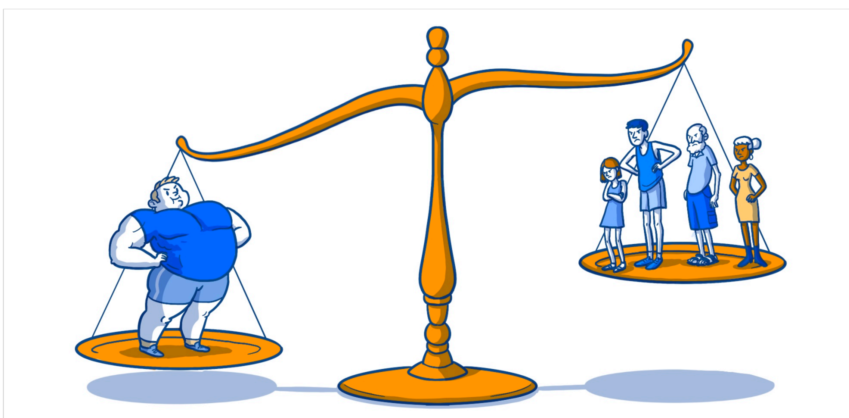
There are many types of bias. People who work with data need to know these so that they can be on the lookout for them.

Selection Bias. This is where data does not contain a balanced range of people. An example would be when we collect feedback on a service for example 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 persons ideas are wrong.



O*utlier 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 planes **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.

Bias Example 1: Kidney data.

Most people are born with two kidneys. The kidneys help to clean waste products out of the blood turning them into urine. At some time in their lives everyone will have a blood test to check how well their kidneys are working. The test measures waste products in the blood and uses a mathematical formula to turn that into a measurement doctors can use to decide if a patient's kidneys are healthy or unhealthy.

We know that some kidneys naturally work better than others but also that some parts of our lifestyle affect how well they work. One of the main waste products in the blood comes from our muscles and this means that this waste product (creatinine) will be higher in people who are older, male and heavier. The amount of muscle you have also affects the creatinine level, more muscular people have higher levels of creatinine in the blood.

Researchers found that 'on average' black people were more muscular than other races. That made them think that the levels of creatinine in the blood could be higher for black people than other races without that meaning there was a problem with their kidneys.

After 20 years this thinking was increasingly questioned. It was found that Black people are NOT more muscular 'on average' than other races. This meant that if a black person was healthy their creatinine level should be the same as the result for other races. If a black person had an increased creatinine level, this was a sign of kidney problems not of being more muscular.

Unfortunately, this mistake has meant that many black people will have had early kidney disease that was not detected and waited longer for treatment than other races. Black people will have been more ill before they could have a kidney transplant than other races and this will have harmed some people.

This mistake was made because the research data did not have a sample that stood for the entire population correctly. That lead to the error of thinking black people had higher creatinine levels due to being more muscular when that was not the case for most black people in the entire population.

This is an example of actual harm from mistakes being made with data collection. But it also has a positive message, a call to action.

We need people of all backgrounds to take part in research and share their data. That is the only way we can make sure no one is accidentally harmed when they receive healthcare.

Bias Example 2: Measuring oxygen levels Our last example showed us that researchers and doctors thought race affected kidney function but in fact it did not. This next example is the opposite. Researchers and doctors thought that a machine to test the amount of oxygen in the blood worked the same in people of all races. We now know that it does not.

Pulse oximetry is a way of measuring the amount of oxygen in the blood. It uses a clip which is put over the fingertip and is connected to a machine which uses light waves to measure the level of oxygen in the blood. Often it is put on one hand whilst a nurse or doctor checks your blood pressure on the other arm. The technology in pulse oximeters was developed many years ago and has been used worldwide to help care for patients. Doctors and nurses thought that the technology worked as well in patients of all races. There was no convincing evidence that showed it did not work the same way for everyone.

But a team has looked at this very carefully. A group of doctors in America studied large numbers of patients and found something that had been missed. They found it because of the large amount of patient data they collected.

They collected 48,097 results from 10001 patients. When they analysed this big sample, they were able to see a problem that had been missed in research on smaller numbers of patients. They found that the pulse oximeter machines did not work as well in black people compared to white people. The machines 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 be missed by the technology compared to a white person. 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 quickly as they should have.

> *"Black patients were three times more likely to have a low oxygen level be missed by the technology compared to a white person."*

Researchers and doctors thought that people's skin colour did not affect how the technology worked. This problem happened because the first testing of the technology was not in racially diverse groups. Looking at small groups of people and patients the effect that skin colour had on the accuracy of the measurement could not be seen. It needed data from 10,000, with 48,000 results to show there was a problem.

This finding is very significant. Pulse oximetry is used in routine medical care. For patients who are well without breathing problems the problem with pulse oximetry is unlikely to cause problems. But we now have to be more careful with how we interpret the results of that test in patients who are unwell or at risk of becoming unwell with low oxygen levels as we might not detect a problem as quickly as we should.

WHAT DOES THIS TELL US?

These two examples show how we can make mistakes due to the data we use in research not standing for the whole population. It also shows why you often need large amounts of data if you are going to generate safe, reliable results.

The problems with the kidney test were because researchers tried to allow for the fact that there are some differences in how the human body works between different races. But they 'saw' a difference' in their data that was not real. There was no difference in how muscular black people are and so there is no difference in how their kidneys work. The pulse oximetry problem was also because of a data problem but this time the reverse. The researchers had data on too few people with not enough different types of people represented in the data.

This meant they assumed the technology worked equally well for all races when in fact there were differences between races.

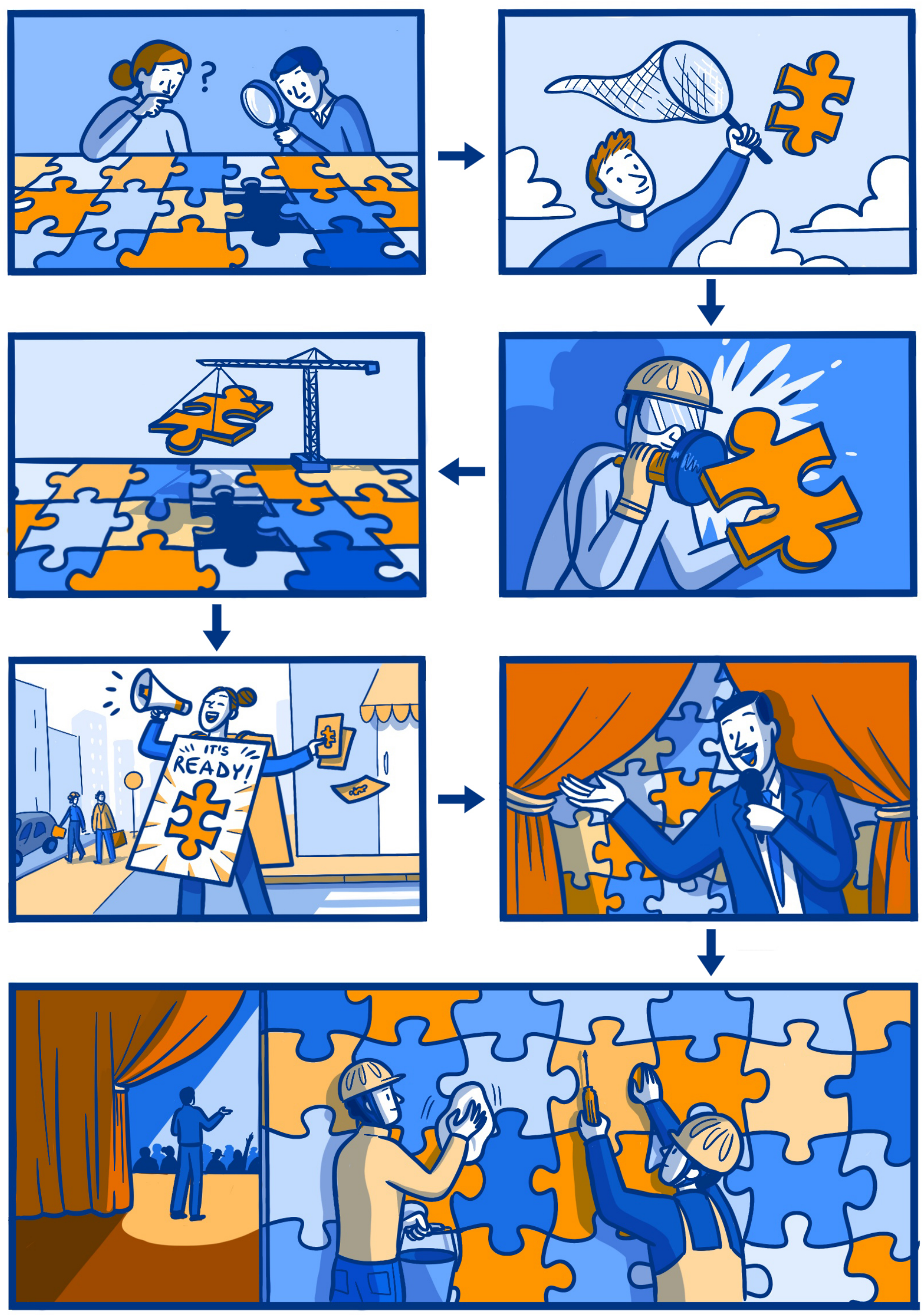
Data being used in research and healthcare technology needs to be correct, complete and reliable and have good representation of the people in the population who will be affected by the research or technology.

CHAPTER 4:A DATA PROJECT/DATA VISUALISATION

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 will explore what we can do to try and make sure the data we are using is high quality. How do we use that data in research or to help us design new services or devices?

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 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 which shows us how what we have learnt in the previous chapters is used in real life to solve problems in healthcare and other areas of the world around us.

A data-science project usually has 7 stages



1. Problem Analysis

2. Data Collection

3. Data Preparation

4. Exploratory Data Analysis

5. Data Modelling

6. Visualisation and Communication

7. Deployment and Maintenance

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 (section 1) or it could involve something called 'scraping' where data is taken from places that are not neat and tidy databases 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 section 3 for example missing and duplicate values. 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 or height in meters and feet.
- Then we need to ask ourselves, does the data seem sensible? Does it look like we would expect?
- 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 Kg, but some were in stones, so we change those to Kg in the data.

✓ **Are any outliers?**

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

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

We check the 60Kg as this is heavier than the other children of the same age. The result is correct. The child is very tall and muscular and wants to take part in the world's strongest women when she's old enough.

✓ **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 2 pieces of data being linked (correlated) and one thing changing another (causation).

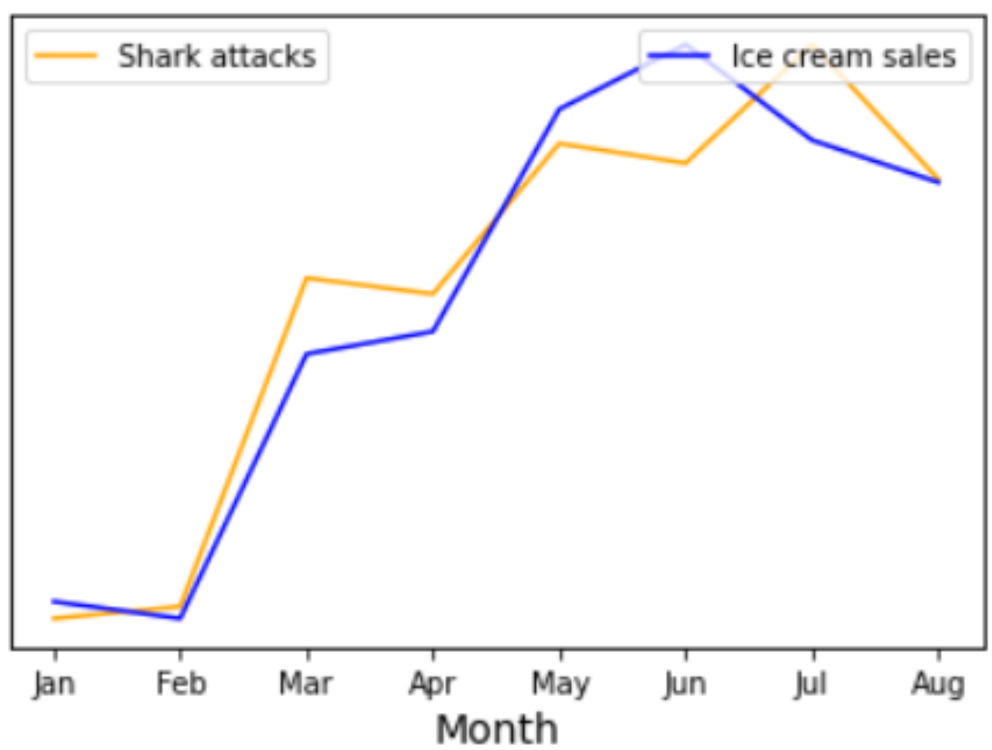
This is one of the most important things in this chapter. Lots of people get this wrong – especially newspapers and that's why we get some attention grabbing headlines that don't actually fit the facts!

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!***

You might be tempted to read the article.



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 are the same. They are correlated.

But that is different from one causing the other.

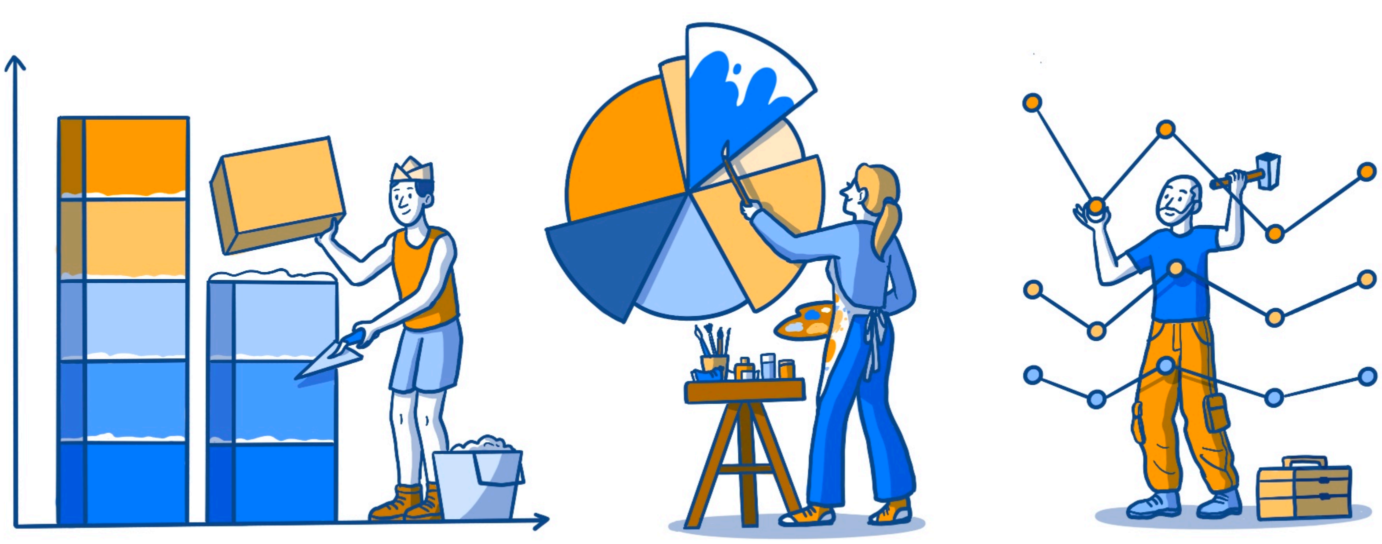
Could eating an ice cream increase the chances of you being attacked by a shark? Could being attacked by a shark make you more likely to eat ice cream?

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 linked to each other. Although ice cream sales and shark attacks are highly correlated, one does not cause the other.

5 **Data Modelling.** The data scientist will run different mathematical tests on the data to find out which help 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 patients risk of a heart attack from data on their blood test results?

The aim of this stage it to be able 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.



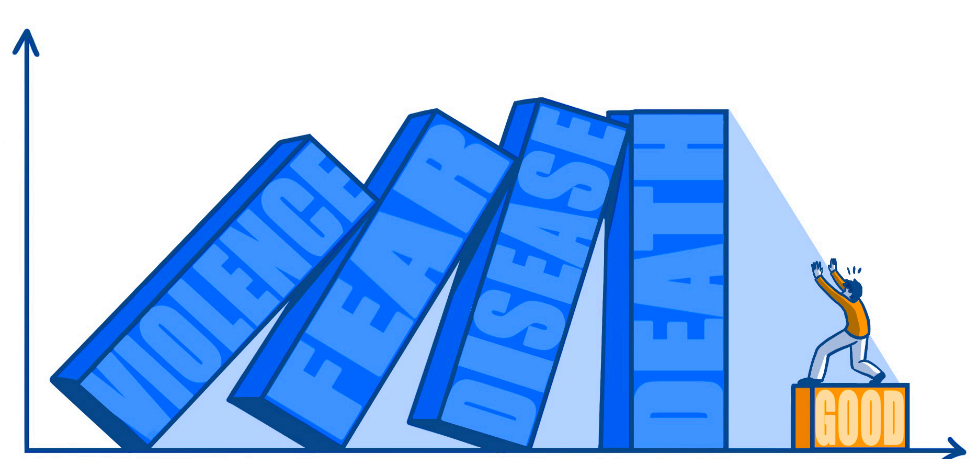
The words we use when we talk about data can affect what people think

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 or 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'.

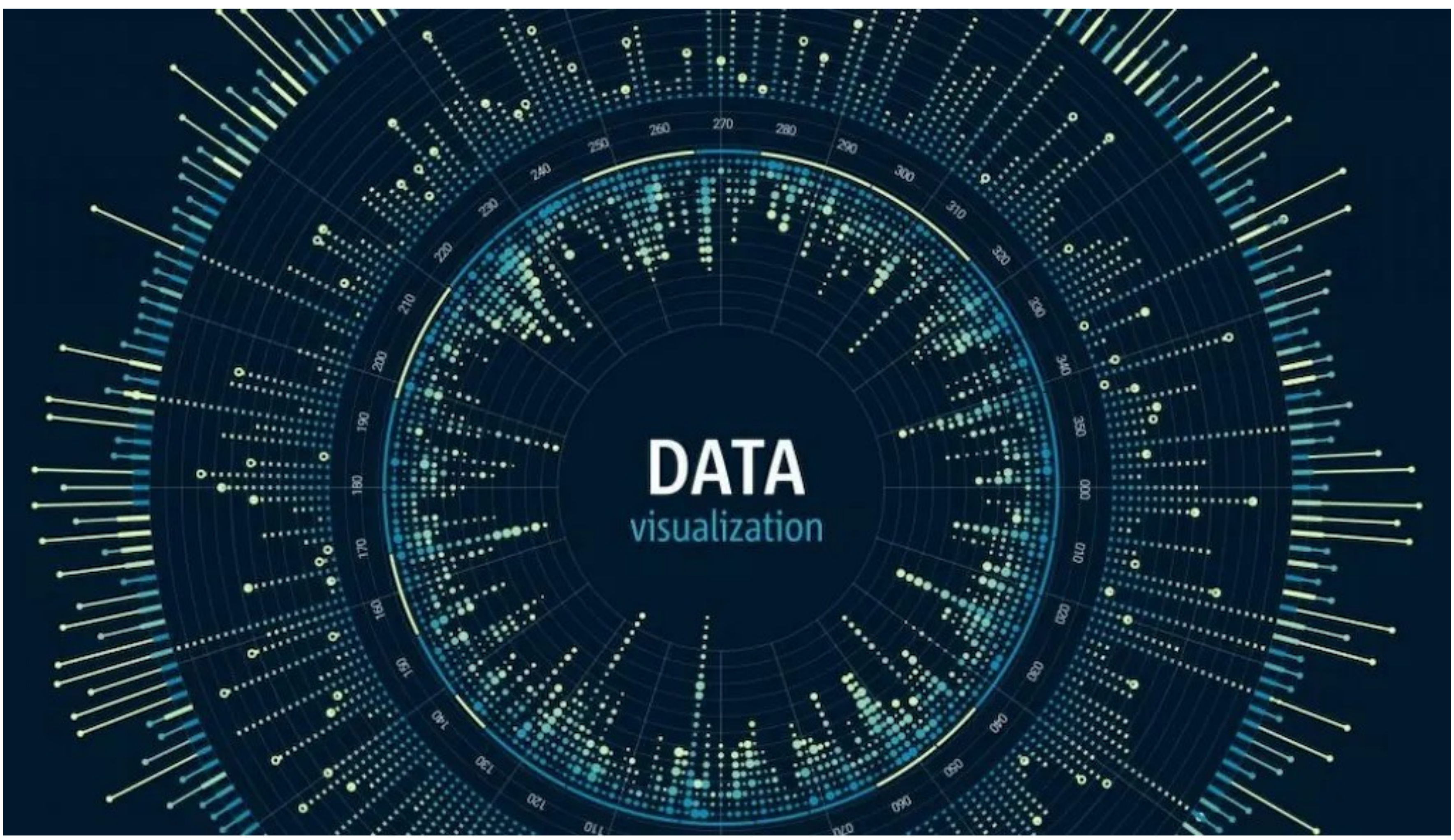
If we wanted to worry people even more, we could use the power of big numbers. Most people don't find it very easy to imagine percentages especially when the percentage relates to a really big number. If we convert the percentage into number it might make it grab people's attention.

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.



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: HOW DATA IS USED IN THE WORLD AROUND US

So far we have thought about what data is and the problem 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 and 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 really powerful so we all need to understand how it can affect us for good and for bad.

Case Study 1: Using data to compare hospitals.

The Bristol Children's Heart Surgery investigation

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. It had been alleged that more children died after being treated there than at other hospitals.

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

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 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? After all these must have been ill children as they needed open heart surgery.

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. This again took a lot of work and understanding to clean and tidy the data so that it could be used in the investigation.

The investigation concluded that we should have 'expected' that there would have been 32 deaths in Bristol. These deaths would be what you would expect to see given 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 be have been avoidable deaths.

So how do you decide what the truth is? Which data source do you believe to be 'more' correct than another?

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 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 closed and some of the staff were found guilty of misconduct.

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.

If you want to read a more in depth discussion of this case study we recommend 'The art of statistics, Learning from data' by David Spiegelhalter.

Case Study 2: Another problem with correlation and causation

Going to university increases your risk of having a brain tumour. That's a headline that will make a lot of people read the article. But what are the facts?

Swedish researchers found a 19% relative increase in the risk of a brain tumour in people who went to university compared to people who left education after school. Sounds like a good reason to not go to university, never mind the student loan you will have to pay back.

What do you think about this? What questions are you starting to ask yourself? Hopefully, you are starting to realise that just because two things follow the same pattern it does not mean that one is causing the other.

Maybe going to university is correlated with an increased risk of a brain tumour because the thing that is 'causing' the increase in brain tumours is something that is also linked to going to university.

But before we get carried away. What do we need to do? We need to look at the numbers and decide how big the effect of going to university is on developing a brain tumour?

We have been told about a 19% increase in relative risk. Think back to the processed meat story. What did we learn? Relative risk is the difference in risk between two groups compared to each other. In this case it is the risk of getting a brain tumour in people who went to university compared to those who left education after school. We are told there are 19% more brain tumours in people who go to university.

What is our next immediate step? The next VITAL thing to do is to ask – 'How many people get a brain tumour who leave education after finishing school? What is the baseline risk of getting a brain tumour?'

Approximately 5 people out of 3000 who don't carry on in education get a brain tumour.

Now we are getting somewhere. Our next step is to work out what is a 19% increase in 5 in 3000? It is 6 in 3000. We now know that approximately 6 in 3000 people who go to university develop a brain tumour. 1 more person per 3000 who go to university develop a brain tumour compared to people who do not go to university. This sounds a lot less worrying than a 19% increase. In fact, this “increase” is so small that it probably is not a difference at all. It is a small difference due to chance.

Case Study 3: Be careful with risk. Is the headline misleading you?

'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.

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

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 a lot of people who might be affected.

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.

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 our 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 life style this difference of 1 in 100 feels different to an 18% increase.

Podcast

This podcast is a conversation between Phil Booth, a data expert, and Jonathan Gregory, a surgeon, around sharing your data. More specifically how to decide what to share and who to share it with. To listen open the camera on your phone and aim it to the QR code. A link will appear on your screen follow that link to the podcast on YouTube.

SCAN ME

CHAPTER 6: AN INTRODUCTION TO ARTIFICIAL INTELLIGENCE

In the last 5 chapters we have been learning about data. Now we will learn about artificial Intelligence. 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 AI is what puts the search findings 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.

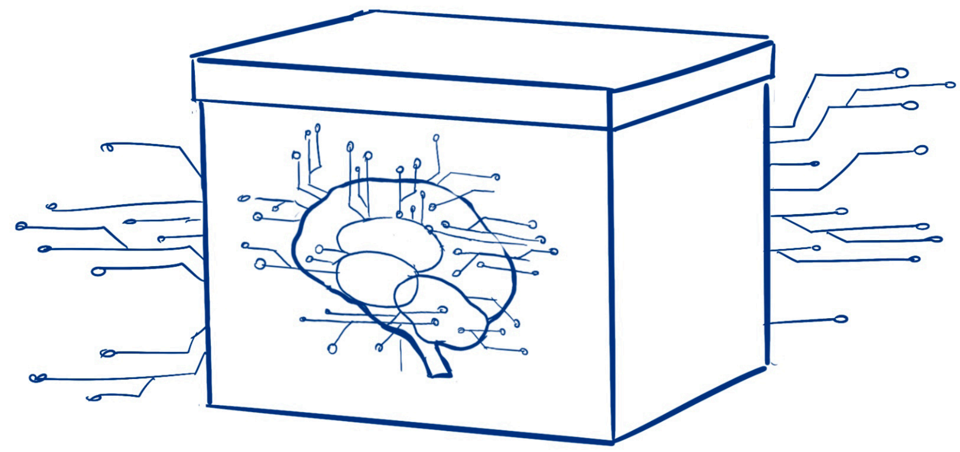
AI is not one thing; it is a range of different computer programs. If we want to explain AI in one sentence we would say 'it is computer technology that does tasks that in the past would have needed human intelligence'. For example translation software. In the past if we were abroad and we wanted to 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 Dany Ruta, an expert in AI, around how translation software uses AI. To listen open the camera on your phone and aim it to the QR code. A link will appear on your screen follow that link to the podcast on YouTube.

SCAN ME

How does AI work?



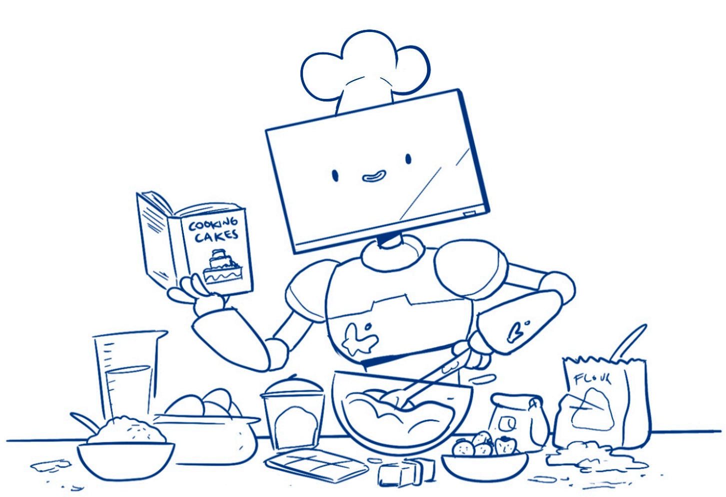
AI is a set of instructions which are written in a computer program (software). The instructions run a computer programme which performs mathematical tests on data. Which mathematical tests used and what the result looks like depends on what problem we wanted the AI to solve. Humans then review the results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

There are lots of different types of AI. Different mathematical test can be used and there are different types of data.

The instruction 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.

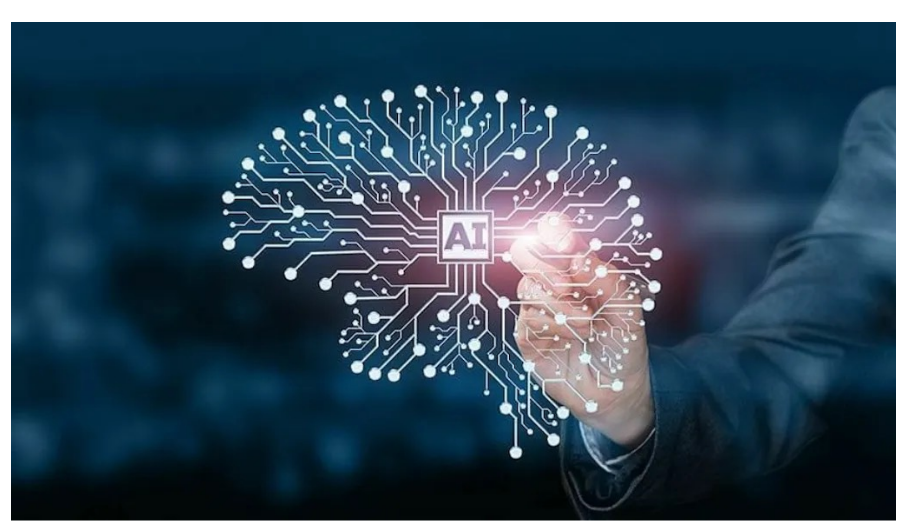
One of the most amazing things about humans is that we are so good at so many different things, like language, mathematics or using common sense when things don't go to plan. The list of our human abilities is enormous. There is not a single AI system that can do all of the things that a human can do.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

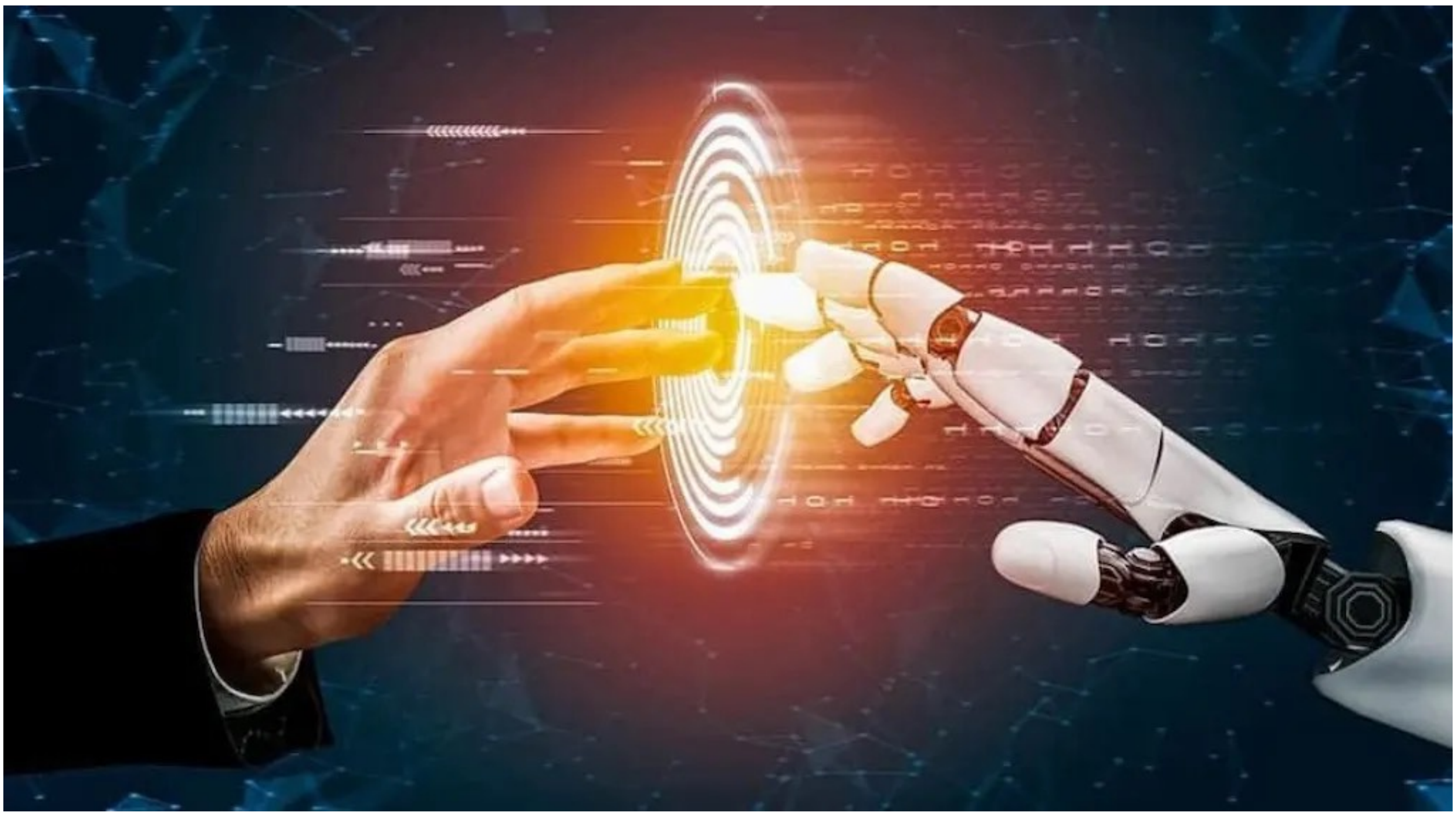
A computer follows a programme that allows it to solve problems using mathematics. That programme contains an algorithm, and it is this algorithm that gives the computer 'artificial intelligence'.

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 put the English dictionary onto a computer and then ask for the definition of a word. 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 give it information on what treatments are used for different sizes and types of breast cancer. We can then give the AI details of patient's cancer and ask what treatments the guidelines say they could have. 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 made.

Now medical science has become so advanced 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 rules that they use are created by humans. The computer algorithms that are created with these rules must be programmed into the computer by humans. For example there may be guidelines on what treatment should be used with a particular test result and that is what the human programmer will put into the expert system AI. 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.



Algorithms made by computers – machine learning

Humans can write a computer programme 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) to the computer. The computer looks for patterns in the data that take it to an answer to the problem it has been asked to solve. Humans provide feedback on the answers the AI generates and these feedback into the AI and it will adjust the mathematical formulas to get closer to solving the problem. Gradually it turns these patterns into mathematical formulas that allow it to solve questions that we ask it. It may find ways to group pieces of data together or to make predictions from data.

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 1000's 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 or ways of finding cancer in blood test results.

If we used machine learning, we would give the computer blood test results including people who we knew had cancer and people who didn't. The machine learning algorithm could be asked to predict which had cancer and which didn't. Humans would check the result, and give some corrections where the AI was wrong. The computer would then change 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 used these to make predictions.

Are computers really learning?

When we start talking about computers 'learning', that's when some people start to get worried; scary science fiction movies spring to mind!

Machines or computers don't really 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.

An algorithm is a recipe. It is the rules an AI follows from data. Humans can make the recipe and give it to a computer. These are the rule based systems. The computer can make the recipe itself by testing thousands of different recipes and finding the one which works best.

This is machine learning.

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!

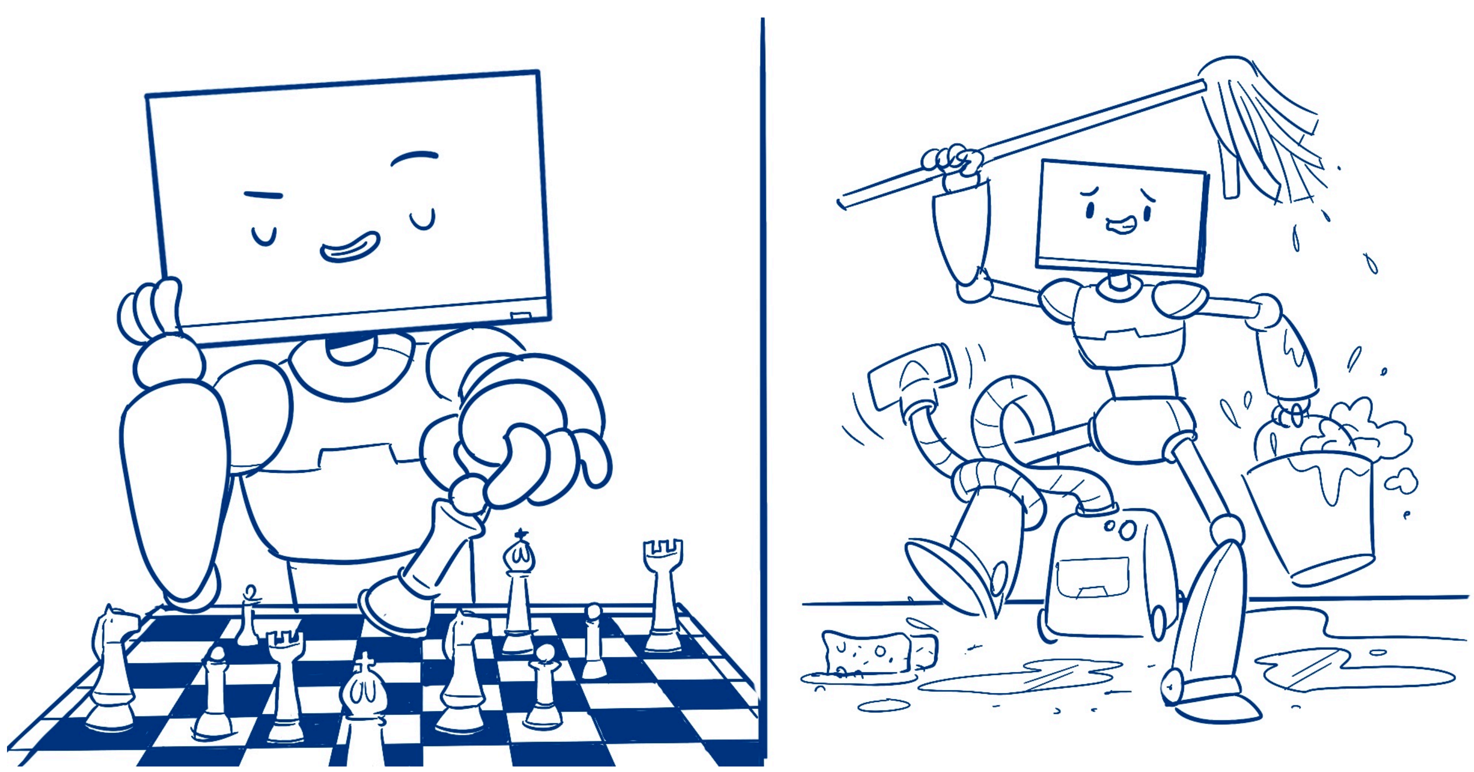
A company called Deep Mind developed a machine learning AI called AlphaGo which beat the Go World Champion Lee Sedol. The scientists that programmed the computer did not teach the computer how to play, they just gave it the rules of the game. The algorithm then evolved by playing thousands of matches against amateur 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 AI 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.



	Rules 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	A chatbot on a website which asks questions and gives responses	A computer system that finds patients who might have a heart attach 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: MACHINE LEARNING AI IN HEALTHCARE. THE PROBLEMS.

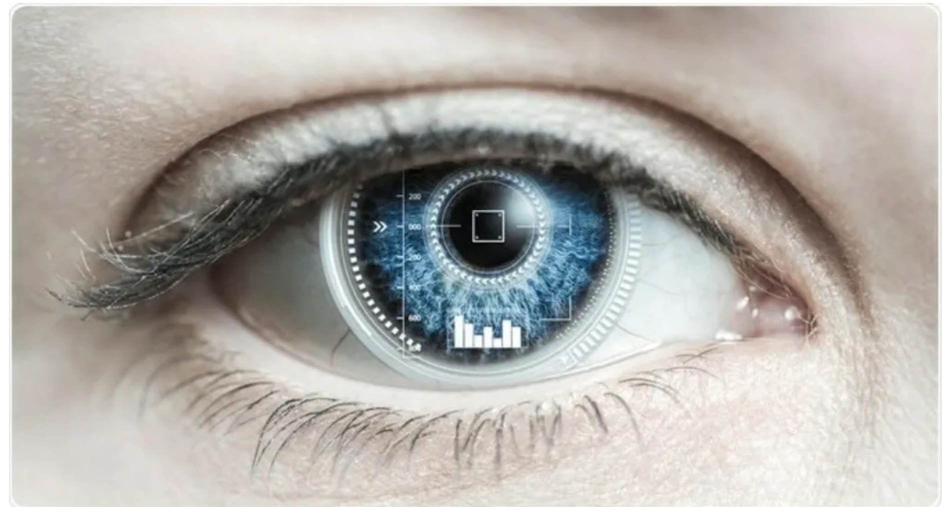
If we use an AI that is based on rules that humans make, then it is easier to check and understand if it is safe. Being sure machine learning AI is safe is much more difficult.

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 that rules 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.

Google have created a machine learning of 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. The AI can find from the retinal photos 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 so the retina photo AI didn't work as well in all places. It has needed lots of work and testing so it can be used safely in different countries.

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

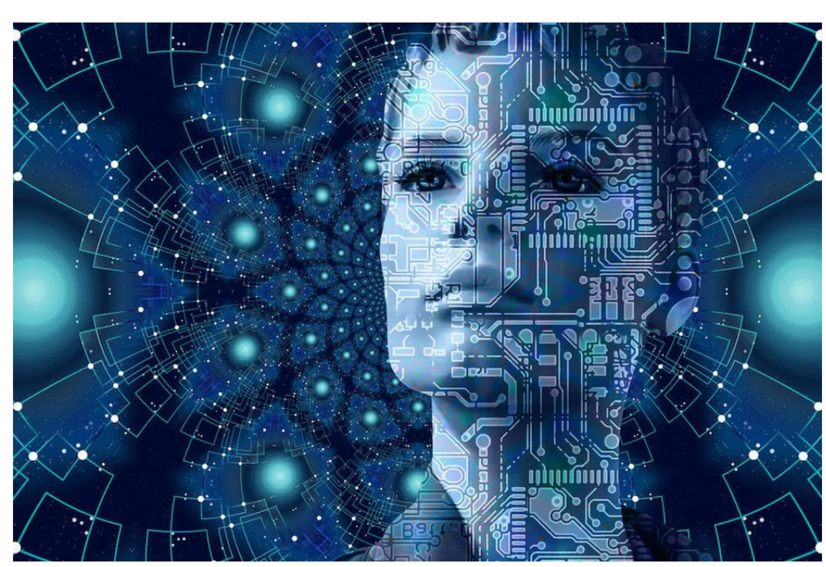
Generalisability is how well an AI works when you use it in a different place from where it was made. When you use the AI somewhere else it will be using data that is slightly different from the data that was used to make it. The data might be different because it involves different types of people for 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 could be slightly different.

Covid 19 and AI

At the start of the pandemic hospitals were overloaded with patients and tests for covid sometimes took several hours to get a result. 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 AI's was biased. This meant that that the algorithms did not work well in different hospitals.

Another problem with machine learning AI is that we don't always know what algorithms the AI is using so we can't explain why the AI gave the answer it did. This is called **Explainability**.

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 means we can 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 and why it chose to do it that way.



Scientists are working on ways to help us understand what the machine learning AI has done but these are not yet perfect.

Bad data makes bad AI

Let us revisit what we read 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 a bad AI. A bad AI can mean it doesn't work very well, it gives us the wrong answers, or it 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 help predict which patients with serious chest infections (pneumonia) could be sent home safely and who needed to be admitted. The data showed that people who also had asthma were less likely to die from pneumonia. This was spotted by the computer. In the algorithm it used this data to make it more likely to suggest sending people home who had asthma.

But there was a problem with the data. The data did not show that because doctors were very worried about patients with asthma who had a serious chest infection, they gave them 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 something about what it is to be human. When an AI doesn't work as we expect but we learn something really important. In this case if we treated some patients with serious chest infections who don't have asthma with more intensive treatment would we get better outcomes?

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 the AI was being made it will have had 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.

Spotting breast cancer on x-rays

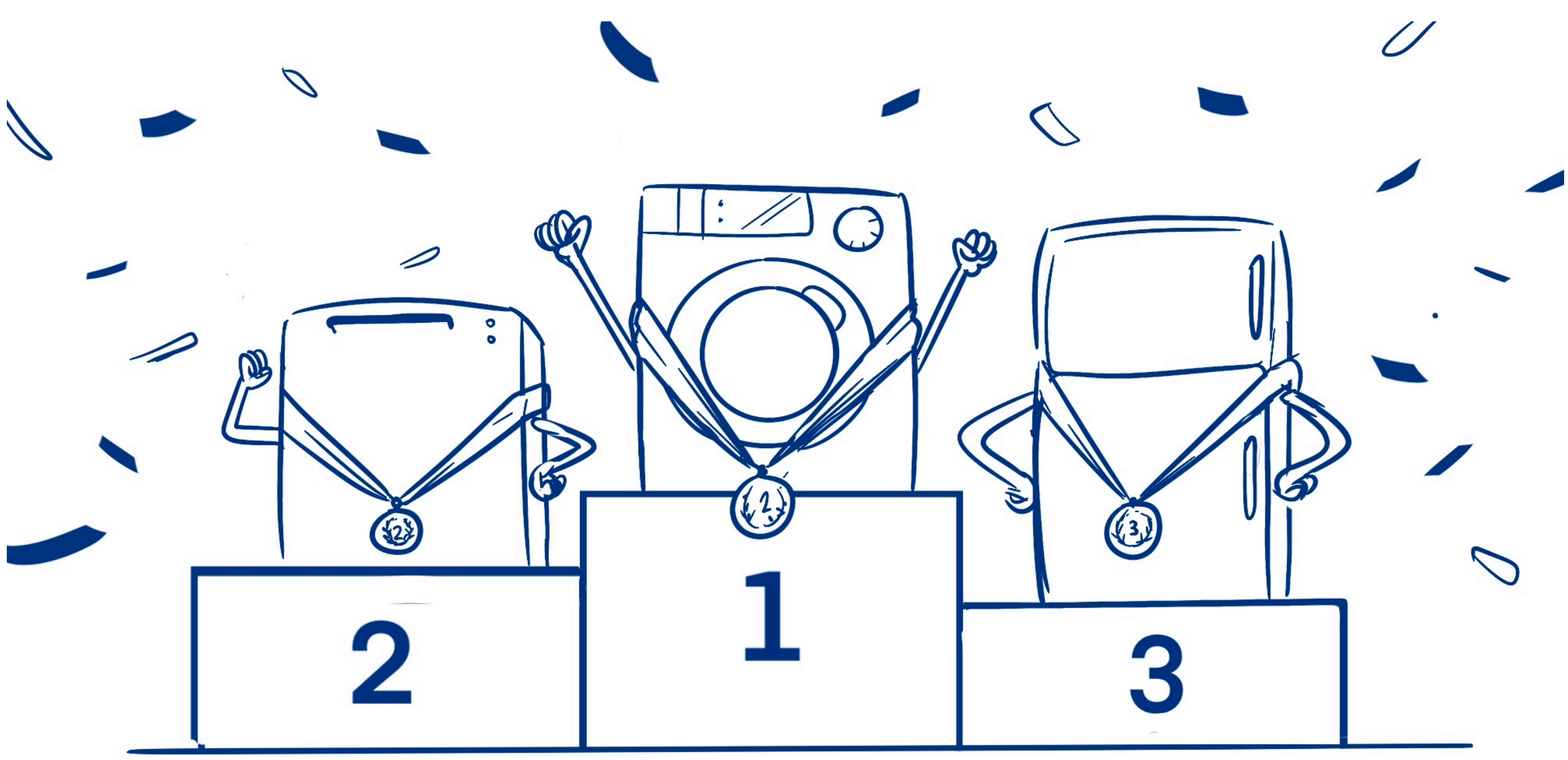
There are lots of people trying to create AI systems that can help diagnose breast cancer from x-rays of the breast (mammograms). But human breasts are not all the same. Some have more fat in them others have more fibrous tissue which makes them denser. Women from all backgrounds can have different types 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 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 these women. This would mean that it might not work as well for women of Chinese ethnicity. The same problem could happen in reverse if an AI was made 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 anything 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. Rules 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. Because of this, 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 (see the Examples, of how AI might improve healthcare table below). 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 MACHINE LEARNING 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 having a 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 and with no human intervention or training. 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 noticed.

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

Podcast

This podcast is a conversation between Jonathan Gregory, a surgeon, and Dany Ruta, an AI expert, around the use of AI in healthcare. To listen open the camera on your phone and aim it to the QR code. A link will appear on your screen follow that link to the podcast on YouTube.

SCAN ME

Examples of how AI might improve healthcare

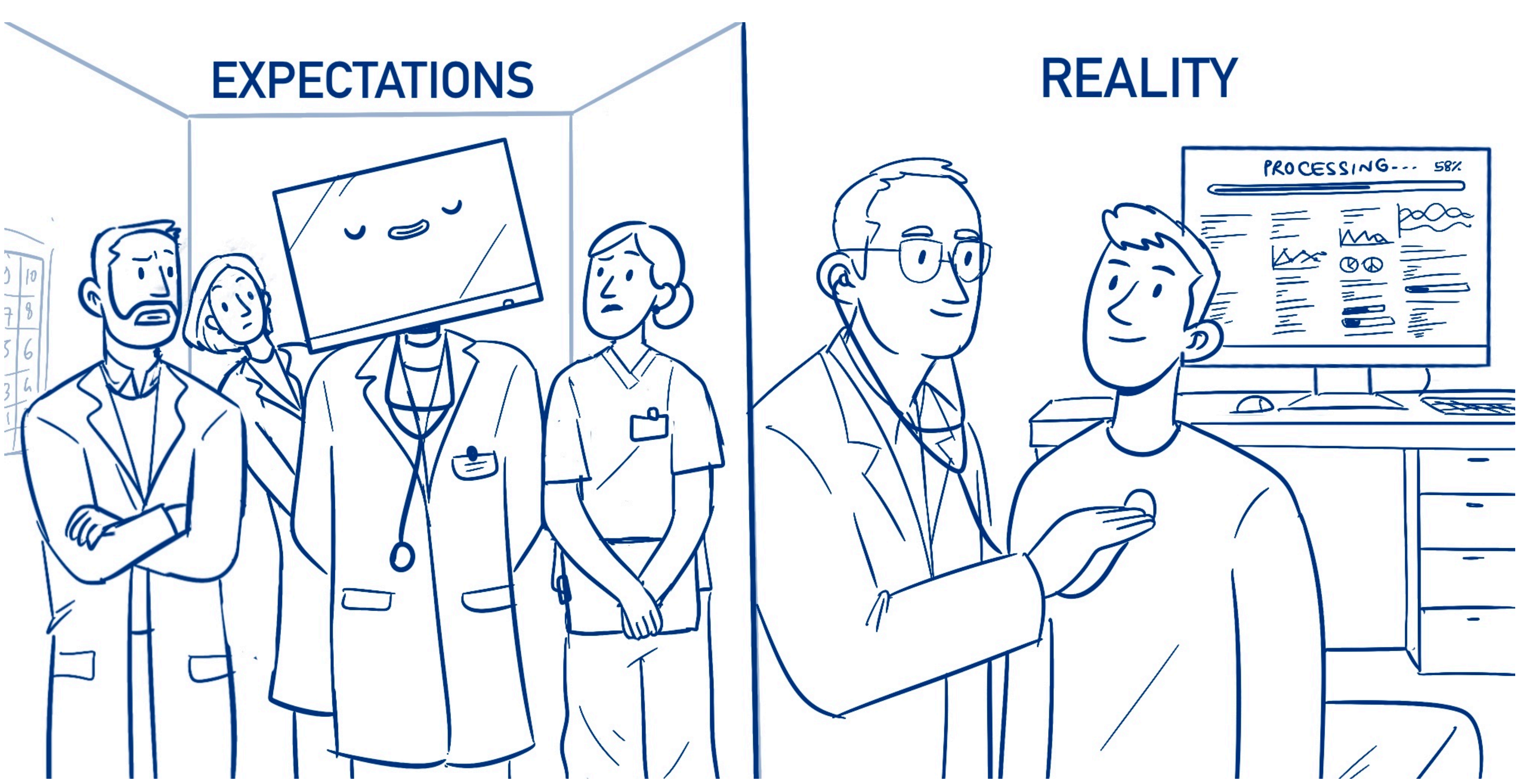
Problem	How might AI help healthcare teams?
There are not enough x-ray specialist doctors, so patients wait a long time for scan results.	AI might be able to check all the x-rays and scans and find those that are normal. It could create an automatic result that was sent to the patient and their doctor 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 test, it could trigger humans to double check if the AI things something is abnormal, but the humans thought it was normal.
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.
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 will alert the nurses to problems but if everything is OK the nurses can 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 help with staff shortages by AI helping to prioritise what is important.

When AI is going to be used to help healthcare staff decide about 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. At the same time looking at the patient to build a relationship with them and reading their face and body language get clues to what the problem might be. 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 was 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 patients' 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 teams we 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 have 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, physiotherapist is harder now than at any other time due to the huge advances in medical science. There is now too much to know for any human to keep up to date. 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.

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 new technology can be made which works for everyone without bias?

ACKNOWLEDGMENTS

Public Contributors

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

Subject Matter Experts

Hatim Abdulhussain
Katherine Baker
Zach Bobbitt
Phil Booth
Dorota Chapko
Alan Davies
Ang Davies
Ronnie Davies
Ilianda Eleftheriou
Clive Flashman
Kelly Gleason
Jonathan Gregory
Glen Hart
Francis Hooley
Charlie Huins
Chris Kelly
Suzanne Kelly
Halle Johnson
Kerlann LeCalvez
Julia Letts
Rachita Mallya
Martha Martin
Kelly Mousa
Kabelo Murray
Hina Pandya
Maria Piggin
Stepahine Posavec
Danny Ruta
Poitr Teodorowski
Vishaal Virani
Matt Williams

We would like to thank our funders.

The 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 – Cancer Theme