A podcast on how to understand research data,

with Professor Michael Harris

This podcast provides an overview of statistical terms and research methods relevant to primary care.

02:51

Introduction

Being able to **interpret data** is really important to us as GP's and I'll tell you why. To give good care to patients, we need to seek out the relevant **clinical evidence** and then evaluate it. To be able to evaluate a piece of evidence, which may be a research paper or an article, we need to be able to find out which **research method** has been used and why. We have to know the basics of how that research method works.

For **quantitative research** (that's research with **numbers**), we need to be able to interpret the results, which means knowing how to understand the statistics that **describe** those results.

We don't need to know how to do the statistics, but we need to know how to understand what they mean. Then, we need to decide whether or not to change our clinical **practice** in view of that clinical **evidence**.

There is a huge range of different research methods and statistical techniques, and in this podcast, I'll go through the important ones and explain when they're used. There isn't nearly enough time to explain all of them in detail. This podcast, in effect, gives you a signpost to the relevant terms.

04:03

Medical research

Medical research is about answering a **research question**. When we see a patient, we may ask ourselves what's the best way to look after this patient? Is treatment **A** better or worse than treatment **B?** What do patients withs this condition worry about? How can we best help them?

Medical researchers translate our clinical question into a **research question** and that's what research papers try to answer. They pose a research question, they explain how they're answering it and they tell us what the answer is. So, if we're looking at a research paper, the first thing to do is find the research question. Is it **relevant** to our clinical practice? If it is, is it a question about how or what people **think**? Or is it about **numbers**? If it's about what people think, it's called **qualitative research**. If it involves numbers, it's **quantitative research**, so let's discuss that first.

05:02

Quantitative Research

Quantitative research answers research questions by collecting numbers known as **numerical data** and then analyses those numbers with statistics. There are different types of quantitative research. So, let's go through which ones are used and when.

Let's say we're interested in whether **marathon runners** eat **apples** and the effect that it has on their marathon **performance**.

Many thanks to Professor Gordon Taylor, a statistician from Exeter University, for these completely fictitious examples.

05:35

Descriptive studies

We could start with a descriptive study. We could survey a sample of marathon runners and ask them how many apples they ate in the last month, and we would use **descriptive statistics** to describe the results.

What **percentage** ate any apples? What was the **average** number that they ate? This could be given as a **mean** for evenly distributed data or a **median** for skewed data, with an indication of the spread of those answers using **standard deviation** or SD for short for **means**, or **interquartile range** for **medians**.

We could find out about the **incidence** of eating apples (how many started eating them in the last year) and the **prevalence** (how many are eating them today).

06:45

Confidence interval, Chi-squared test and *P* value

If we found out that in a **sample** of 100 marathon runners, 30% of them eat apples, how near might that be to the percentage of **all marathon runners across the country** who ate apples? A statistician would calculate something called a **confidence interval** to tell us that, so we need to be able to interpret that.

Then we may want to know whether there is a significant difference between the number of men and women who are eating apples. We could use a **Chi-squared test** for this.

The ***P* value** is the end result of this and a lot of other statistical calculations, so we need to be able to interpret it.

In this case, ***P* value** from the Chi-squared test will tell us if there is really **no difference** between the groups, how **likely** is it that we'd have got the difference we found in this group. If the result was unlikely to have happened by chance, with the ***P* value** of **less than 0.05**, we say this is **statistically significant**.

It's a really important concept. If we see that a ***P* value <0.05**, it means there probably really **is** a difference between the two groups.

08:04

Correlation, *r* value

We can look for an **association**, known as a **correlation** between the number of apples that runners eat and their marathon times. Do marathon runners who eat apples tend to have **faster** marathon times, **slower** times, or is there **no pattern**?

Researchers use an ***r* value** as a measure of that, so we need to know how to interpret an *r* value.

All this means is that there's an association. It *doesn't* give evidence of **causality**. It's not evidence that one causes the other. It may be that there's a **confounding factor** - that there's another underlying cause for this association. For example, marathon runners who eat apples may be more or less likely to drink more coffee, and that's the real reason for the link with their marathon times.

09:08

Case-control study, odds ratio

What if we are interested in a **rare** outcome? For example, what's the association between regularly eating apples and actually winning marathons? Then, we would need to do a **case-control study**.

Cases are the runners who have won a marathon.

Controls are runners who have not won a marathon.

The **predictor** we're interested in is self-reported regular consumption of apples.

We would use an **odd ratio** to compare the two groups.

09:46

Cohort study, risk ratio

Let's say that the preliminary results from our cross-sectional and case control studies suggest there is a link between eating apples and improved marathon times - that's an association, it hasn't shown a causal link. The next step is to do a prospective **cohort study**. This is a comparative, observational study, in which the marathon runners are grouped by their

exposure status (whether or not they eat apples) and then follows them up overtime to see what their marathon times are.

Cohort studies start with the **exposure**, not the **outcome** and we use a **risk ratio** to compare the results for the two groups. It gives some evidence as to causality. That it is the exposure (which is eating apples in this case) that has caused a difference in marathon times.

As with correlation, there may be confounding factors that explain the findings. This is an example of what is known as **bias**, which is a **systematic error** in the study that affects the results. I've already mentioned coffee intake as a possible confounding factor. Another possibility is that people who eat apples may be more or less likely to do pre-race stretching exercises and it's that difference in whether or not they stretch before the race that causes the difference in the marathon times.

11:14

Randomised controlled trials

So now we need to consider doing a **randomised controlled trial** (**RCT** for short). If we randomise marathon runners to either eating apples daily or let's say, eating oranges as a control, we can compare the marathon times. The **randomisation** to either eating apples or eating oranges reduces risk of **bias**. If we're evaluating **RCTs**, we need to know about concepts such as **blinding**, **cluster randomised controlled trials** and **parallel group randomised control trials**.

The likelihood of real differences between the two groups is measured with a ***P* value**.

12:15

Type 1 and type 2 errors

When you are looking at ***P* values** in randomised controlled trials and other quantitative studies, we need to be aware of two possible types of error.

One is called a **Type 1 error**. Think of it as a **false positive** – apples don't really have any effects, but the study erroneously suggests that it did. If ***P*<0.05**, that means there's a 1 in 20 chance this was a false positive or type 1 error.

The other is called a **Type 2 error**, which is a **false negative**. Think of it as a **detection failure** - apples really do have an effect, but our study didn't find it, possibly because the sample size was too small.

13:08

Survival analysis, Kaplan-Meier method

Perhaps you want to compare how many months or years it takes new runners in each group to start winning marathons. The authors may use a **survival analysis** technique. A common method for survival analysis is the **Kaplan-Meier method**, to measure any differences over **time** between the two groups.

13:34

Meta-analysis, forest plot, funnel plot

Perhaps a few other researchers have already published papers on this. A **systematic review** identifies the evidence by summarising the healthcare literature through a full literature search and critical appraisal of individual studies.

Have several studies all examined the same numerical research question? If so, we may find a **meta-analysis** – a statistical synthesis of the numerical results to get an overall view of what the research shows. If we look at a meta-analysis paper, we need to be able to interpret a **forest plot** and a **funnel plot**.

14:25

Risk

Running a marathon is really tough and not everyone can complete them. Let's say researchers found that eating apples reduces your **risk** of failing to finish a marathon race. How can we describe that **change** in risk? There's a variety of ways in which we can describe a change in risk of something happening. They are commonly used, so we need to be aware of these.

Absolute risk reduction, **relative risk reduction**, **number needed to treat** or **number needed to harm** and **Cates plots** are the most important ways to describe that **change in risk**. So, as GP's, we need to know what they mean when we see them.

15:22

Diagnostic testing

Another topic that we need to know about to be able to interpret evidence is **diagnostic testing**. How good, for example, is a new diagnostic test for finding or excluding an illness? For that we need to be able to understand and calculate **sensitivity**, **specificity, predictive values** and **likelihood ratios**. These are relatively simple calculations that you can do yourself.

15:58

Qualitative research

We've been talking about the concepts that you need to understand for quantitative research. Equally important for GPs is being able to understand and evaluate **qualitative research**, research about what people think. Think let's say that we want to find out what marathon runners **think** about using apples to help them with preparation for marathons or why they use them. That needs a qualitative approach. We can get data on this through **interviews**, **focus groups** or **surveys**.

Researchers can use more than one of these methods. For example, using **open-ended questions** in a survey as well as interviews, and see how the results from these methods compare; that's called **triangulation**.

16:53

Sampling methods

For qualitative studies, we need to have an idea about different possible **methods of sampling** from all the marathon runners out there. Be aware of qualitative sampling methods called **stratified**, **convenience**, **maximum variation** and **snowball sampling**.

17:26

Data saturation, thematic analysis, mixed methods studies

Another bit of qualitative research jargon that you need to be aware of is **data saturation**, where we keep collecting data until we're not getting any new information. This means that there's no minimum or maximum number of interviews or participants. We also need to know about a common quantitative analysis method called **thematic analysis**. In primary care, we often see what's called **mixed methods studies**. This is where researchers **combine** qualitative and quantitative approaches in a **single** study.

18:13

Validity

Finally, when we've read and understood the results of any study, we need to decide whether they are likely to be **valid**. That means how well do results from the study participants represent **true findings** among similar individuals outside the study? Do the results represent the truth in the population we're studying and are not due to methodological errors? This is called **internal validity**.

Then we need to decide whether they're applicable to other groups of patients, not just the ones that the researchers have selected. That's called **external validity**, or **generalisability**. For example, a study on marathon runners in a very hot country may not be generalisable to runners in the UK.