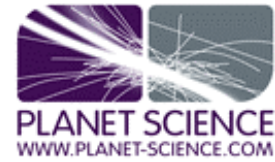




Scary Medical Statistics: Lost In Interpretation!

Teacher's Notes



Scientists don't often like to talk about their research to journalists, as they're really paranoid it will be misrepresented...and in some cases they're right! Often it's because data is wrongly interpreted. Medical scare stories, are usually just that....stories! Some stories could even give Stephen King a run for his money!

This set of activities is slightly different from the rest, since it's all to do with underlying mathematics in science stories and the problems associated with data interpretation.

There are several ways this lesson could be run: it could be a follow on from a practical where you collected some data, so you could use this lesson to look at ways of interpreting the data. Students could also work out ways to "invent" data especially for this lesson, and then analyse and interpret it. Ideas are given below. This needn't be just one lesson, perhaps a couple, or joint with maths teachers (as mentioned in the "Adaptations" below).

Aim:

To show that not all the statistics and medical data we read about in newspapers are as trustworthy as they might first seem.

Learning objectives and outcomes:

- To critically read a piece of literature and analyse the use of science and statistics.
- To recognise that medical scare stories are often blown out of proportion and why.
- All students will understand that data can be interpreted in lots of ways.
- Most students will understand the impact of this misinterpretation on society.
- Some students will be able to apply what they've learnt in their maths lessons to this lesson.

Resources required:

Examples of statistics in the media, i.e. recent news footage where numbers have been used (e.g. percentages of patients suffering from a particular disease), student worksheet, different coloured beads or packets of sweets (e.g. Smarties - "*Only Smarties have the answer!*"), packs of playing cards, calculators (is mental arithmetic no longer fashionable?).

Safety:

If using the sweets to generate data, and you are in a lab, make sure the students aren't tempted to eat them! Otherwise this activity should take place in a classroom.

Starter activities:

Why not start with "Interactive Odds". This will get the students thinking about risk and probabilities. Initially the students will be using their intuition as to the likelihood of each event, but then they'll see the numerical probabilities for each event. This could be done as a whole class activity, followed by a short discussion on why they decided something was more likely than another, and whether they were right? Would it have been easier had the numbers been there from the start?

Introduce the idea that statistics can often be interpreted, and conveyed, in more than one way...some of them are misleading.

A student worksheet is provided, so the students could work through this in the lesson and for homework. The number generator could also be used as a whole class participation activity.

Core level main activity:

1. Start with a quote from Dr Ben's Bad Science article to illustrate different ways of looking at one set of figures:

"Let's say the risk of having a heart attack in your 50s is 50% higher if you have high cholesterol: that sounds pretty bad. Let's say the extra risk of having a heart attack if you have high cholesterol is only 2%. That sounds OK to me. But they're both talking about the same (hypothetical) figures. Out of a hundred men in their 50s with normal cholesterol, four will be expected to have a heart attack; whereas out of 100 men with high cholesterol, six will be expected to have a heart attack. That's two extra heart attacks. Those are natural frequencies. Easy."

Explain to the students that it is often this conversion of natural frequencies into percentages, ratios or probabilities, which makes the original scientific data either more confusing, or completely wrong.

Mention the examples given in the article, and how natural frequencies easily communicate risk. What do the students think about this? Do percentages, probabilities and ratios help them understand a news story better?

2. As mentioned in the worksheet, ask the students to work in pairs or in small groups. Give them either a pack of cards or a packet of sweets (sweets of many different colours are best). And ask the students to generate some data. They can do this by using simple probabilities e.g. on average, how many blue sweets are there in a packet (a good example if they're using the sweets, not the pack of cards!)? They will need to know total numbers of sweets in packets, take averages, and also look at how many blue sweets there are per packet on average. And so encourage the scientific process within this short data collection exercise.
3. They would also show an example of bad data collection i.e. just look in one packet, how many blue sweets there are out of a total; so for example if they have four blue sweets in a packet of twenty, they could state: "For every packet of sweets, you will have four blue ones" or "20% of sweets in each packet are blue" etc. You can highlight the importance of survey size and random sampling at this stage.
4. Alternatively you can provide some data (from a real scientific paper, if possible, e.g. from <http://www.pubmed.gov>, or see data on blood glucose levels provided with this activity and "Adaptations" section below), and ask the students to analyse the figures in a way that it is accurate and a way in which it no longer says what the researcher said.
5. The way the students present their data (good science style and bad science style) can be in the form of a table, graph, pie-chart – any suitable method.
6. Then ask the students (either within the class or for homework) to write a short newspaper article either "sensationalising" the science statistics, or being a really good, trustworthy scientific journalist! Or how about a

PowerPoint presentation? The students could even take the badly interpreted data from another pair/group and pretend that they are Dr Ben Goldacre writing about some newly discovered Bad Science!

Plenary:

Find out what the students thought about data collection and interpretation. Discuss whether the press deliberately or accidentally manipulate the data: is it a matter of inaccuracy, incorrect simplification ("dumbing down"), not fully understanding the data or just wanting to get the best story?

Take-home message: If you're reading a medical/health scare story, you should be asking: am I being given the actual evidence and *who* is at risk?

Extension:

As mentioned in the Bad Science Article, even the legal profession gets in a muddle over numbers and probabilities.

For those of you feeling brave, and wanting to tackle a "crime scene investigation"-type maths problem, look no further than the prosecutor's fallacy (or even the defendant's, interrogator's, or jury observation fallacies). DNA evidence has been used for many years to help convict the guilty and free the innocent. However there are examples where a simple misunderstanding of conditional probabilities leads to innocent people being locked up! Be warned, bringing this up in a discussion with a class is not for the faint-hearted. There is still much debate over these fallacies and how to prevent mathematics getting in the way of proving someone's innocence. The same misunderstanding has also led to investigations where there has been more than one cot death in a family (why should more than one cot death automatically mean that it's suddenly a suspicious case, rather than thinking that the probability is the same each time?). If you're feeling really, really brave, how about adding a dash of Bayesian probability...

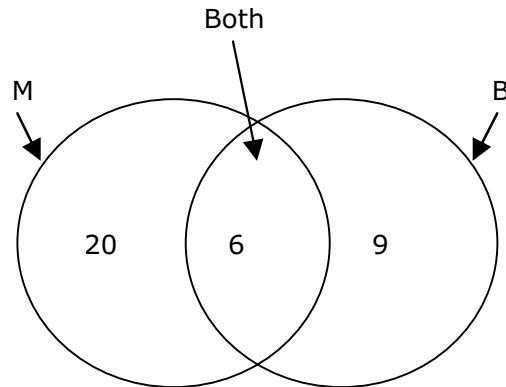
Some useful websites on the topic include:

- Prosecutor's Fallacy: <http://www.colchsfc.ac.uk/maths/dna/discuss.htm>
- Evaluating Legal Evidence (DNA samples): <http://www.dcs.qmul.ac.uk/researchgp/spotlight/legal.html>
- The "Jury Observation Fallacy": http://www.dcs.qmul.ac.uk/researchgp/spotlight/jury_fallacy.pdf

For those of you who would rather just talk about conditional probability (who wouldn't?) and a way in which data is misinterpreted, how's this for a close-to-home example?:

Maths and Biology: Who's studying what?

There are 35 students in the year group (yes, it's a small school!). The number of students studying maths (M) is 26, taking biology (B) is 15, and six students study both. Students can visualise this better through a simple Venn diagram (as shown below).



The question to ask: So, if you were to walk into the biology classroom and pick a student at random, what is the probability that they also study maths?

The common mistake here is that people assume that the conditional probability of students doing maths given that they do biology is the same as the conditional probability of students doing biology given that they do maths. Not the case, at all. So how do we explain this?

Using the Venn diagram above, if we went into the biology classroom there is a total of 15 students ($6 + 9$). Six of these students also study maths. Therefore the probability of picking a student *doing maths given that they study biology* is $6/15$. If you were to walk into the maths classroom, there is a total of 26 students ($6 + 20$). Six of these students also study biology. So the probability of picking a student *doing biology given that they study maths* is $6/26$.

So we can conclude that there is a higher probability of a biology student also studying maths than there is of a maths student also studying biology.

Adaptations and Other ideas:

- Since there is a limited amount of mathematics you can cover during a science lesson, perhaps this could be a joint lesson with the maths teacher (or run the activities in both maths and science lessons in the same week)
- Get the students to act out a radio/television interview (and actually record it) where the interviewer introduces the scientist having got the research completely wrong. The scientist then comes out, and the discussion continues with the scientist making his/her case about the real findings of the research.
- The background information sheet provided with this activity could be a useful starting point to have other discussions...especially on the topic of gambling (how people are attracted into playing the lottery even though there is a 1 in 14,000,000 chance of winning (a six number lottery)!
- If you feel limited by time, or wish to set homework or an extension activity, why not incorporate a *real* medical situation into the lesson, such as diabetes, and the need to monitor blood glucose levels. The data provided (blood glucose levels) can be analysed by the students, instead of getting the

students to generate their own data! The actual data was used to look at homeostasis, the mechanisms that regulate blood glucose levels in the body.

Information to provide with blood glucose data:

Your blood glucose level is measured in millimoles per litre (mmol/l). During the day your blood glucose level is between 4 and 8 mmol/l, but they are higher after meals and usually at their lowest level when you wake up in the morning (which is why breakfast is so important!). In diabetes, the blood glucose levels move outside of this normal range.

Ideal blood glucose levels should be:

- 4 to 8 mmol/l before meals
- Less than 10 mmol/l one and half hours after a meal
- Approximately 8 mmol/l at bedtime

ICT:

Looking at websites where there has been misrepresentation of statistics and frequencies in science topics in the media.

Keywords:

Statistics, probability, frequency, fallacy, data, mathematics.

Citizenship:

Significance of the media in society.

Assessment opportunities:

Scientific literacy skills, critical thinking, mathematics in science, data collation and interpretation, data presentation methods (graphs, pie-charts).

Curriculum links:

This can be incorporated into lessons where you may be looking into the importance of testing explanations, the need for scientific literacy and numeracy, and considering the strength of the evidence.

Homework tasks:

Did you ever believe the media over a health story and then find out it was flawed research, or the data had been misinterpreted? What was it, and what made you believe them? Was there ever a case where it seemed obvious that the media had glossed over the facts and made assumptions on *very little* data?

Useful links:

- BBC KS3 Bitesize:
http://www.bbc.co.uk/schools/ks3bitesize/maths/handling_data/probability1_intro.shtml
- Checking out "the facts and figures behind the news" (in the U.S.):
<http://www.stats.org/index.jsp>
- Royal Statistical Society Centre for Statistical Education:
<http://www.rsscse.org.uk/>
- Experiments at School: <http://experimentsatschool.ntu.ac.uk/>
- Other statistical lesson ideas: http://www.stats4schools.gov.uk/lesson_ideas/

Medical Scares!

Everyone would like to think that drinking and eating what you like is good for you...but aren't we just manipulating scientific results to fit our lifestyles, and make us all feel happy? If a scientist manipulated their data, technically known as "fiddling the figures", they'd be shunned by the rest of their community! One of the reasons scientists continually communicate their work through conferences and journals is so that others working in their field can point out any mistakes and assumptions they've made with their data.

Here are some other examples of Dr Ben Goldacre's Bad Science articles which could be used as a basis for other useful discussions:

- *A wee dram cuts obesity risk? It's not that simple.*
(<http://www.guardian.co.uk/life/badscience/story/0,,1669542,00.html>). In time for Christmas, a bunch of stories were published about chocolate and alcohol being good for you. Typical! And also how people who drink alcohol in moderation have lower levels of obesity than heavy drinkers and those who don't drink at all...Now, now before we get carried away, is this really what the research showed? Quite a few variables weren't taken into consideration at all.
- *After feeding the scare he'll sell you the solution.*
(<http://www.guardian.co.uk/life/badscience/story/0,12980,1604200,00.html>) . What happens when someone claiming to be a good scientist, but not really doing proper research, is the one everyone believes? Dr Chris Malyszewicz claims that all hospitals are unclean, letting MRSA flourish in each and every corner! But can we trust his data?