Embedding Evidence in Commissioning

It is a fundamental principle of good public services that decisions are made on the basis of strong evidence and what we know works.

What Works: evidence centres for social policy, UK Gov, March 2013

NHS Statutory Duty in respect of research

Each clinical commissioning group must, in the exercise of its functions, promote—
(a) research on matters relevant to the health service, and
(b) the use in the health service of evidence obtained from research.

Section 14Y Health and Social Care Act 2012

Evidence based decisions are those:

- which are taken on the basis of a full analysis of available evidence
- which are monitored and evaluated throughout implementation to ensure that they continue to deliver the desired outcomes

NECS is supported by its Research and Evidence team to provide its customers services, advice and support which are transparently evidence based.

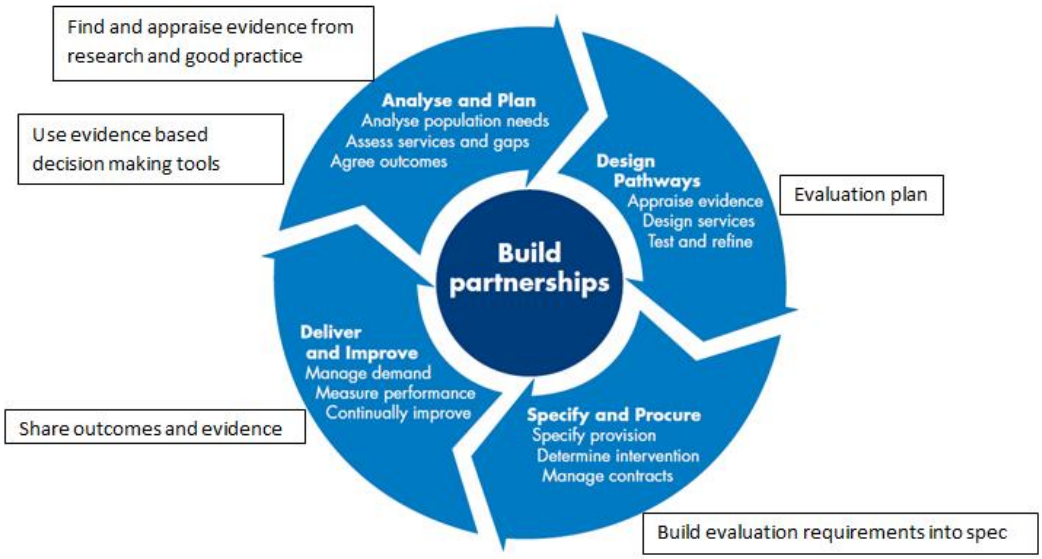
Evidence based healthcare is not new. The whole of modern medical practice is based on observing which interventions were successful and trying to understand why they succeeded.

If the NHS is to provide the best possible care decisions should be taken based on the evidence of what works. In those cases where there is little or no evidence we need to be clear on what we are doing, why we are doing it and how we can know that we have achieved the desired outcome.

The Research and Evidence intranet page brings together resources that will help colleagues to find, evaluate or create the evidence to support. To ask for more details, training or support from the R&E team mail NECSU.RETeam@nhs.net

Using the evidence base and incorporating evidence gathering and creation into the commissioning cycle are the best ways of ensuring that commissioners get value for money. Commissioners will have a firm understanding of what works and delivers the desired outcomes, and they can be assured that new ideas and services will be monitored and evaluated as they are implemented.

What evidence is will often depend on context; there may be a lot of evidence for something, or a little; evidence may be in one place or dispersed amongst a number of sources. For this reason it is important that NECS staff members are able to assemble and critique the evidence base when supporting CCGs in their commissioning role.



Adapted from

RC GP Centre for Commissioning

The Commissioning Cycle

Different aspects of the research and evidence function will be needed throughout the commissioning cycle.

Evidence from existing services should be reported to Commissioners to alert them to potential problems or identify areas of best performance which may be replicable in other service areas.

When Commissioners identify areas for service reform or introducing new services, research work is needed to find evidence of what does or does not work in the areas under consideration.

Where evidence does not exist work will be needed to identify what information will be needed from the new or reformed service to assess its effectiveness

All of the evidence researched needs to be critically appraised so that Commissioners can take properly evidence-based decisions.

It is difficult to assess how much research and appraisal work might be needed in commissioning support work. The evidence function will be needed throughout the life of the project but the effort will not be evenly spaced.

There are a number of variables which need to be considered:

1. Is there already national guidance or established best practice in the service area under consideration?
   - Is the existing evidence of high quality? Systematic reviews and Randomised Controlled Trials may not need as much study as small scale, uncontrolled trials or anecdotal evidence.
   - If a new or reformed service is being considered it may take time to define the outcomes.

2. Taking into account the variables above, a good starting point would be to assume 0.5 working days on research and evidence for every week of the project. This is not a hard and fast rule but should be increased or decreased depending on the complexity of the project. The research and evidence work will be needed throughout the life of the project.
3. The research and appraisal work is not an overhead to projects but integral to them.

NECS staff should aim to become involved in commissioning intentions at as early a stage as possible. Ideally, NECS should be able to tell commissioners how services in their area are performing against the outcomes defined in the service specification.

Summary

To ensure the research function is fully involved in commissioning support, consider the following steps:

- When discussions with CCGs begin over commissioning intentions, NECS senior managers should outline the NECS "research and evidence offer"

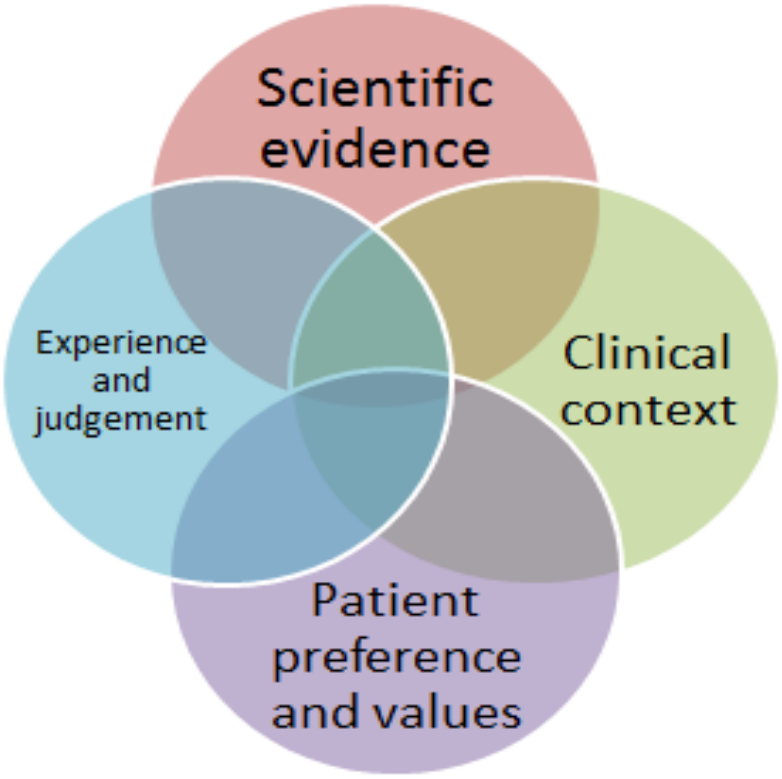
- NECS staff who will be doing the research should sit in on early discussions on commissioning intentions to understand what commissioners want to achieve

- There should be a report back stage to commissioners on initial research; Commissioners decide how best to proceed

Hierarchy of Evidence

Evidence (n) – Information indicating whether a belief or proposition is true or valid
Concise Oxford English Dictionary 12th Edition

Evidence is not mechanistic, that is that it is all about facts and that the facts only point in one direction. This is not so. Evidence will depend on the context it was created and in which it will be used. In any area, there may be lots of evidence or there may be little or none. Even where there is considerable evidence in a particular area, it may not be relevant to the problem you are looking at. Similarly, evidence from a different field of study may have applicability for something you are doing despite the differences. In addition to all of this evidence can be strong or weak. This is why there is a need to look at the wider range of available information.
For the health service, the wider context means any scientific evidence, the experience and judgment of healthcare professionals and the views and preferences of patients.



All of the evidence may be overwhelmingly in favour or against something. Often, though, the evidence is only indicative; that is, it leans in one direction or another. In these sorts of cases, the evidence needs to be interpreted to decide what it is telling us in those circumstances. Where the factual evidence is particularly strong, there may be little need for any other information and vice-versa. As the quality of evidence gets stronger, the more it can be relied on. The diagram below shows the hierarchy of evidence, with the strongest type of evidence at the apex of the pyramid.

Finding the Evidence

OpenAthens is an identity management system which allows access to protected medical databases.

It is recommended that all NECS staff should have an OpenAthens logon.

Contact NECSU.RETeam@nhs.net for more information and support in obtaining at OpenAthens logon.

Any decision to be taken starts with a problem, real or potential, to be solved.

The first step to a correct solution is to define the problem properly. Though this might take time and seem unnecessary and pedantic, it will save effort and ensure that any solution devised is more likely to be the right one.

Commissioning intentions can often be expressed in high level terms. Sometimes this can be because Commissioners want to generate ideas. Sometimes it is because the Commissioners are not sure what may need to be done or how to do it.

The evidence search strategy you will adopt will depend on what you are being asked to produce: generating ideas will need a different search strategy to a need to answer a specific question.

There is no one approved method of defining a question. You should use the method that you feel will give you the best results. You can always change and refine your search terms as you go along.

Make a note of any search terms that you use so that you can easily find again any useful or interesting information that you did not save earlier.

- How has the problem been stated?
- Define the terms used
- Make sure you understand the true nature of the problem
- What is the end product of this query?

When you are confident that you know what is needed you can start to look for the evidence. There are a number of search tools around which will help you develop your search strategy.

a) Try framing the question as if it were the title or an article or essay.
For example, you are asked to look for evidence on community education for type 2 diabetes. You could frame this as: "For adults over 18 does community education for type 2 diabetes result in better self-management (for example, adherence to treatment or dietary regimes) than classroom based education for type 2 diabetes?"

b) PICO – Population Intervention Comparison Outcome
PICO gets you to break your question into its constituent parts which then become your search terms.

Population	Adults 18 or over with type 2 diabetes
Intervention	Community based education
Comparison	Classroom based education
Outcomes	Better self-management, fewer emergency admissions

c) CLIP – Client Location Improvement Professional

CLIP can be used where the search is towards service management.

Client	Adults 18 or over with type 2 diabetes
Location	In the community
Improvement	Patients manage their condition better, fewer emergency admissions
Professional	People appropriately qualified to deliver the training

d) ECLIPSE – Expectation Client Location Impact Professionals Service

ECLIPSE is a more detailed version of CLIP. Like CLIP is aimed more towards health service management. The Expectation element is similar to Improvement in CLIP.

Expectation	Patients will be better able and willing to self-manage their condition
Client	Adults 18 or over with type 2 diabetes
Location	Community
Impact	Improved adherence rates to treatment and dietary regimes, fewer emergency admissions
Professionals	People appropriately qualified to deliver the training
Services	Education services for people with type 2 diabetes

There are many websites which you can search for evidence. Most of these websites require that you register before you can use them.

For health service searches the best place to start is the NICE Evidence database, which contains evidence from a wide variety of sources. The NICE Evidence database can be accessed directly here NICE or through OpenAthens.

Summary - Looking for the evidence

- Make sure your search terms derive from the definition of the problem
- Use the most appropriate search tools
- Don't get distracted

Analysing the Evidence

Appraising the evidence means to assess its value and quality. Before that can be done the data needs to be analysed to understand what it is saying.

Good evidence can sometimes be used to back up inadequate arguments and good arguments can sometimes be supported by inadequate evidence. If we are to make the best use of evidence we need to work out what it is saying and we need to find the evidence that will best support the best ideas.

This page lists a few things to look for when considering from whatever source.

Ask away
The information you are looking at will have been collected for a reason by a particular means. Make sure to understand why the information was collected and what questions were asked to gather it. Information can be skewed due to how and why it was gathered.

Confound it
Information is collected to answer specific questions. There are many variables which might affect (or confound) the data collected and therefore give an inaccurate answer. A confounding factor (sometimes called a confounding variable) is a "variable or factor that was not controlled for in the study but can have an influence on the results." (From "Statistics for Dummies" by Deborah Rumsey Ph.D, p274.)

On average
There are three types of average: mean, median and mode.

- The mean is the sum of a group of numbers divided by the count of the numbers in the group. This is the most commonly used and understood term for average.
- The median is the mid-point, that is it has the same number of values on either side of it
- The mode is the value that occurs most often in a group.

Take the group: 1, 3, 4, 4, 4, 5, 7, 8, 9

- The mean is 5 (45/9)
- The median is 4 (there are four digits on either side)
- The mode is 4 (4 occurs most often in the list)

Unless it is clear from the context, the type of average being used should be specified.

One, two, three
Three data points are needed before it is possible to talk about a trend. One data point says only something about itself. If there are two numbers, one will either be higher, lower or the same as the other.

Missing you already
Numbers given without context are not very helpful. For example, a statement that "98% of people" support something means little because it is not known many people were asked to express an opinion and of those how many replied. Equally unhelpful are statements which make unsupported assertions; for example, "There is widespread dissatisfaction with service X". Ideally, information should be given with its context but if it cannot, supporting information should be openly available and easily accessible.

All Change
Healthcare commissioning will often look at comparisons, including looking at data over a period of time. If comparisons are used, are they valid? For example, are averages being against aggregate data? If attendances at a clinic are being compared over a number of years, was the method of counting the attendances the same in each of the years? If there have been changes which might affect the comparison, they should be noted and allowance made for them.

It's a probable risk
Probability is an assessment of how likely something is to happen; something is a risk if it will have a bad effect. It is wrong to assume that random occurrences in the past have an affect on random occurrences in the present or future. Something may be highly improbable but that does not mean that is impossible. For example it is improbable that a coin tossed 100 times will come up heads on each occasion but it is not impossible.

When considering the probability of risk, there is a difference between the absolute risk of something happening and the relative risk. The relative risk of something might be quite large but if the absolute risk remains low, the impact of any change might be small.

Significantly confident
Researchers will want to know how likely it is that the results they have observed could have happened by chance.

A significance level will be calculated. Historically, 95% is the accepted level of statistical significance. If the results are statistically significant then it is unlikely the results happened purely by chance.

A confidence interval is a range of values calculated from the data between which the true value lies. The narrower the interval, the more accurate the results are.

Axis of misdirection
When showing information graphically make sure the image does not give a false impression. When using graphs, make sure the values on the axes are helpful and do not give the wrong impression; for example if the scale is too small, small changes in values might look very big. If you are using diagrams or shapes, doubling the height of a column or radius of a circle will result in the area increasing by four times; again this could give a misleading impression about small differences.

Bigger than huge
When dealing with big numbers it might be helpful to reduce the amount down to what it means for each patient, or GP; this will give people a good understanding of what the number represents for the individual.

It sticks out a mile
Outliers are values which are very high or very low. (Though the term is usually used in relation to numbers, it can include qualitative information, for example, someone who has an extremely negative opinion about something when other opinions are broadly positive.) If there are outliers try and find out why they are so different.

The Simpson Paradox
The Simpson Paradox is a situation where a conclusion made from aggregated data is reversed when the data is disaggregated. The problem occurs when a confounding factor is overlooked when aggregating the data. To help avoid this situation look to see that possible confounding factors have been identified and allowed for.

Appraising the Evidence

Evidence (n) – Information indicating whether a belief or proposition is true or valid
Concise Oxford English Dictionary 12th Edition

Evidence does not exist in a vacuum; it has been created for a reason and by particular methods. The reasons for the creation for the evidence, the ways it has been collected and the conclusions drawn from it all have to be understood before it can be said whether or to what extent the evidence properly answers your question. Put more simply, critical appraisal is about assessing whether the evidence is good enough to be included in our decision making.

Critical appraisal is not simply about finding reasons to dismiss research, it is not a purely statistical exercise and it is not something which only experts can do. Critical appraisal skills can be learned by anyone. Once learned they will help you assess the usefulness of any research.

There are many reasons why research and evidence may not be valid:
Bias - Is the evidence biased? Bias occurs when evidence leans more heavily in one direction because it has been collected in way which systematically excludes or over represents some types of data. For example, a door-to-door survey conducted during the working day will exclude anybody who is at work.
Clarity – Is the evidence presented clearly? If you are not able to understand the evidence it may not be your fault. It may be that the evidence is not well presented or explained: for example, there might be too much jargon, assumptions may have been made which are not explained or the evidence presented does not address the issue raised.
Misleading comparisons – if the evidence compares two or more groups, are the groups genuinely comparable? For example a group which has many members may not be comparable with a group which has far fewer members.
Confused outcomes – the evidence may report outcomes that just do not make sense.
Bad maths – the numbers presented in the evidence might not be correctly computed or the wrong numbers might have been used.
Confounders – there may be more than one reason for any results presented in the evidence you are studying. For example, a study shows that excessive smoking causes early death. However, it may be that people who smoke excessively also drink to excess, take little exercise and generally have unhealthy life styles. The early deaths may be caused by a combination of factors not just the smoking.

Critical Appraisal Summary
Critical appraisal is about answering three questions:

- Is the evidence valid?
- What is the evidence saying?
- Does the evidence help locally?

The Critical Appraisal Skills Programme (CASP) has produced a series of checklists which can help you assess different types of evidence

Sources of Bias

- Bias is an inclination for or against someone or something. Bias can affect the way evidence is collected and interpreted.
- Bias in collection occurs when evidence has been collected in a way which systematically excludes or over represents some types of data. For example, a door-to-door survey conducted during the working day will exclude anybody who is at work.
- Bias in interpretation occurs when we believe or disbelieve evidence for irrational reasons. For example, someone dismisses the evidence about the value of a treatment because it does not suit their opinions.
- Bias can be conscious or unconscious. Even if it is conscious, it is not necessarily malicious. People do like to make the strongest case for their point of view and this might affect what evidence they use of how they present it.
- Bias is not something that just affects other people.

If a bias is unchecked it will produce bad results which will lead to bad decision making.

Listed below are some types of bias but there are many more. The best way to check bias is always ask yourself why you believe something.

Anchoring bias
This happens when a person relies too heavily on the first pieces of information they receive. In effect, the first piece of information becomes an unconscious benchmark against which other evidence is measured. One way of avoiding this form of bias is to find other comparative data.

Availability bias
This happens when people are influenced by information which is easily available. This could be because the information is more recent, or perhaps it is something that happened to you and is therefore more memorable. One way of avoiding this form of bias is to find out as much data as possible about the subject you are studying.

Clustering illusion
This is the tendency to see patterns in random distributions. Clusters can occur in random sets of data; the bias comes about when something significant is attached to these clusters. For example, a study in the 1970s suggested that childhood cancer was associated with children living near power lines. In fact, the clusters of childhood cancers were random but the researchers had noted the closeness of power cables to the homes of the children. The researchers should have asked whether the clusters were significant and if so what might be the cause.

This sort of bias is also known as "The Texas Sharpshooter". The nickname derives from the story of the Texan who fired his gun randomly at the side of his barn, then drew a circle round where the bullets were most clustered and claimed that was the target all along.

Confirmation bias
This happens when someone only takes account of information which supports their ideas. For example, someone has already decided on a particular course of action: they research only the evidence which supports their decision and ignore any evidence against it. A variation on this is the Ostrich effect which is the tendency to ignore unwelcome or dangerous information.

Post hoc, ergo propter hoc

This is a Latin phrase meaning "After this, therefore because of this". This is the belief that because B follows A, A must have caused B to happen. For example, if A&E attendances go down after a new service has been introduced it would be wrong to assume that the attendances have gone down because of the new service. You need to consider all the reasons for why A&E attendances may have reduced.

This type of bias is sometimes referred to as "Correlation is not causation".

Sampling bias

This happens when you use a population sample which is not properly representative. Unless you are able to question everyone in a particular group (say the population of your home town) the best way to get an idea of what people are thinking is to question a representative sample of the population. If the sample is big enough and genuinely representative you should get a good idea of what people are thinking. If the sample is not big enough or not genuinely representative the results will not be as helpful and could be useless. In these cases the results are said to be affected by sampling error. Calculating the correct sample size can be quite technical and you may need help from colleagues in Business Intelligence. (There are some sample size calculators available on the internet but they should be regarded as indicative.)

Survivor Bias

This kind of bias can happen when you take a sample of a population now and look at the history of that sample. The mistake can occur because you are taking your sample from the existing population and missing those who did not survive to be included. For example, if you are studying risk factors to the health of people in their sixties you may be excluding important information because you are not looking at those who died in their fifties. This kind of bias can be mitigated by taking your sample population from the past and examining their history forward over the years.

Summary

Bias in the way evidence is collected could result in inaccurate and misleading data. Bias in the way evidence is interpreted can lead to bad decision making.

There are many types of bias. The best way to avoid them is to ask yourself why you believe a particular piece of information.

Statistics

Statistics are a way of representing data numerically and lots of people are wary of them. When used properly statistics should be clear, precise and unambiguous.
In healthcare, one of the main uses of statistics is to try and quantify the difference between two treatments or activities; that is, is one better than the other?

When looking at any statistics ask five questions. (The list is taken from "How to lie with statistics" by Darrell Huff.)

Who says so?
Who has produced the statistics? Are they from a reputable source? Do those presenting the statistics have an interest in the outcome, for example selling a product or promoting their services? Conflicts of interest should be declared so that you can make a proper assessment of the information presented.

How do they know?
What kind of research has been done to collect the data being presented? The results of a well conducted, large scale randomised control trial are likely to have more validity than those of a small scale survey only conducted in one location.

What's missing?
Data often needs context before it can be properly understood; properly presented data should include all of the information you need to understand it correctly.
Has any data been left out? Have all of the participants in a survey been accounted for? Where percentages are used, have the original numbers also been provided?

Did somebody change the subject?
Does the data support the conclusion reached? For example, a survey asks people whether they take their medicine correctly. Most people reply "Yes", leading to a report that most people take their medicine correctly. The data does not support the conclusion because the data refers to reported activity and the conclusion to actual activity.

Does it make sense?
Overall, do the data and the conclusions drawn from them, make sense? If you have to make unwarranted assumptions about the data to agree any conclusions then there is probably something not quite right about the information or conclusions presented.

Uncertainty

- Probable (adj) - likely to happen or be the case
- Probability (n) - the extent to which something is probable

Concise Oxford English Dictionary 12th Edition

Something is said to be uncertain if it is not known whether or not it will happen.

Probability is the degree of chance that the something will happen. The higher the degree of probability, the greater the chance that the something will happen.

If the something that might happen is dangerous then that is a risk.

Life is full of uncertainty. Much of the knowledge the human race has acquired has helped reduce uncertainty but uncertainty has not been eliminated. There are still risks associated with even the most commonplace activities.

If we are to properly understand the efficacy of new treatments we need to understand the benefits and risks associated with them. This will mean that clinicians can correctly inform patients about the risks of any treatment or medicine they may need. The patients can then make properly informed decisions about their healthcare.

The most basic comparison is between taking a treatment and doing nothing (which also includes continuing with any existing treatment). Medical researchers who want to make these comparisons will design suitable tests and present their results for public scrutiny. If the results are valid (validity can be said to be the extent to which a test measures what it was intended to measure) and presented correctly they can be used to determine the extent to which a treatment can have a beneficial effect.

There are two ways of measuring the effectiveness of a medical treatment when compared to others;

- The first way is to measure the Absolute Risk Reduction (ARR) of the treatment. The ARR shows the number of people who experience the outcome of the new treatment compared to the numbers experiencing the outcome from the other treatment measured. For example, a study of people who get headaches looks at two treatments: A (the new treatment) and B (the existing treatment). Treatment B is the control event because we have not changed anything; Treatment A is the experimental event. Calculate a control and event rate by taking the number of reduced headaches in each group divide by the total number people in each group. There are 100 people in each group. The results show that for every 100 people treatment A reduces the number of people with headaches by 6 and treatment B by 4. The ARR is 6% - 4% = 2%.
- The second way is to measure the Relative Risk Reduction (RRR). The ARR will give two numbers: those experiencing the outcome of both treatments studied. The RRR shows the difference between these two numbers as a proportion. Taking the numbers from the previous example, the difference between 6 and 4 is 2; 2 is 33.3% of 6 so the RRR is 33.3%.

One final piece of information needed to help understand uncertainty is the Number Needed to Treat (NNT). This calculation will tell you how many people need to have the new treatment to reach the one additional person who will benefit by it. The NNT is calculated as 1\ARR. From the above example this is 1\0.06 = 16.6 recurring, or 17). Where NNT is concerned decimals are always rounded up.) In our simple example 17 people need to take the new treatment to get one who will experience fewer headaches. Inversely, 16 people will receive a treatment which does not reduce the number of headaches.

Presenting these kinds of results can be problematic. In the above examples, the new treatment does seem to be successful in reducing the number of people who experience headaches. From the point of view of the individual patient, they will be most interested in how likely it is that they will be in the group that does not experience headaches. The patient may feel that the absolute numbers are so small there is little chance they will benefit; on the other they may think the relative chance is one that would benefit them. For the purposes of example, simple numbers have been used but in real life, things will be more complicated making it more difficult to decide whether a new treatment will be beneficial to an individual.

The problem of whether a new treatment will be good for an individual is further complicated because everyone perceives risk differently. Some people do not like taking risks while other people will dare anything. The best solution is to educate people about risk and allow them to make up their own mind.

In considering risk, it is important not to confuse improbability with impossibility. Something may be highly improbable but this does not make it impossible.

Uncertainty – and therefore risk – are part of life. They cannot be entirely eliminated but they can be quantified so that informed decisions can be taken as to whether it is acceptable or desirable to take a risk.

Summary

- Uncertainty is a feature of life
- Uncertainty can be quantified as a probability
- Let individuals decide what probability and risk mean to them

If you want more information, training or support on any of the content,
please contact the Research and Evidence team
NECSU.RETeam@nhs.net