THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Reasoning under uncertainty

Citation for published version:
Aitken, C & Mavridis, D 2019, 'Reasoning under uncertainty' Evidence-based mental health, vol. 22, pp. 44-48. DOI: 10.1136/ebmental-2018-300074

Digital Object Identifier (DOI):
10.1136/ebmental-2018-300074

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Evidence-based mental health

Reasoning under uncertainty

Colin Aitken, PhD[1*], Dimitris Mavridis, PhD[2,3]

Affiliations

1. School of Mathematics, Maxwell Institute and Joseph Bell Centre for Forensic Statistics and Legal Reasoning, The University of Edinburgh, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, U.K.; cgga@ed.ac.uk.

2. Department of Primary Education, University of Ioannina, Panepistimioupoli, Ioannina, 45110, Greece.

3. Epidemiology and Statistics, Sorbonne Paris Cité Research Center, METHODS Team, Paris, France

Abstract: 189

Text:3388

Tables:1

Figures: 1

Correspondence to: Dr Colin Aitken PhD, School of Mathematics, Maxwell Institute and Joseph Bell Centre for Forensic Statistics and Legal Reasoning, The University of Edinburgh, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, U.K.; cgga@ed.ac.uk.

ABSTRACT

Introduction: It is difficult to reason correctly when the information available is uncertain. Reasoning under uncertainty is also known as probabilistic reasoning

Methods: We discuss probabilistic reasoning in the context of a medical diagnosis or prognosis. The information available are symptoms for the diagnosis or diagnosis for the prognosis. We show how probabilities of events are updated in the light of new evidence (conditional probabilities/Bayes theorem). A resolution is explained in which the support of the information for the diagnosis or prognosis is measured by the comparison of two probabilities, a statistic also known as the likelihood ratio.

Results: The Likelihood Ratio is a continuous measure of support that is not subject to the discrete nature of statistical significance where a result is either classified as `significant' or `not significant'. It updates prior beliefs about diagnoses or prognoses in a coherent manner and enables proper consideration of successive pieces of information.

Discussion: Probabilistic reasoning is not innate and relies on good education. Common mistakes include the ‘prosecutor’s fallacy’ and the interpretation of relative measures without consideration of the actual risks of the outcome e.g. interpretation of a likelihood ratio without taking into account the prior odds.

Keywords: prior odds, posterior odds, likelihood ratio, conditional probability, Bayes theorem, prosecutor’s fallacy

INTRODUCTION

There is an implicit assumption in all that follows that uncertainty can be measured by probability. There is a very good argument that probability is the best measure of uncertainty (Lindley, 2014). A clinician is uncertain of the diagnosis, given symptoms displayed by the patient or of the prognosis that follows from the diagnosis. Uncertainty is colloquially presented in statements of likelihood. For example, it may be stated that `it is very likely the patient has a particular disease' or `it is very likely that the patient will die within a certain specified time period'. This paper assumes that such measures of likelihood can be represented numerically by a number between zero and one, a number which is known as a probability. In other words, probability represents a measure of belief.

There is a fundamental theorem underlying reasoning under uncertainty. The theorem is Bayes' theorem, named after a nonconformist theologian, Thomas Bayes (1701 - 1761), who was a student at the University of Edinburgh and a Fellow of the Royal Society. Bayes' contribution to science was two-fold. First, he argued, as above, that uncertainty about the occurrence or otherwise of an event can be represented by a probability. Second, he showed through his theorem how one's uncertainty about the occurrence of an event can be revised in the receipt of evidence or information of relevance to that event. These contributions are applicable to medical reasoning. First, the event, about whose occurrence or not the clinician is uncertain, is a diagnosis or prognosis. The information of relevance to this event is a symptom in relation to a diagnosis or a diagnosis in relation to a prognosis. Two examples show how these ideas work in practice. Throughout these examples, the exact numerical figures are not important for the argument concerning the interpretation of evidence. It is the underlying principle that is important. In both examples, the probability of an event is updated in the light of new evidence. Conditional probabilities are easily misunderstood and prior odds of an event should be taken into account when trying to interpret them. A common problem arising with interpretation of conditional probabilities is known as the *prosecutor's fallacy* (Thompson and Schumann, 1987) from its misuse on criminal cases where a small probability of finding evidence on an innocent person is confused with the probability that a person with whom the evidence is found to be associated is innocent.

This confusion is not specific to criminal cases and this may be further clarified with the following comparison.

If I am a monkey I have two arms and two legs	If I am guilty my DNA matches that of a profile from a crime scene stain	If I have the disease then the test result is positive
If I have two arms and two legs am I then a monkey?	If my DNA profile matches that of a profile from a crime scene stain, am I then guilty?	If the test result is positive do I then have the disease?

In all of these cases, the first statement in each column is true, assuming no false negatives. However, the answer to each of the questions in the second statement in each column is `maybe'.

Example 1: Relationship of a diagnosis to a prognosis

This example concerns the consideration of a prognosis given a diagnosis.

- *Prognosis:* The prognosis is that a person will commit suicide. This is the event about whose occurrence or not the clinician is uncertain. Denote this event as S^+. The event that a person will not commit suicide is denoted S^-.
- *Diagnosis:* This is the information of relevance to the event. The information is that a person has been diagnosed with a psychiatric disorder at least once in their lifetime. Denote this information as A^+.

It is possible to use measures of belief such as personal opinion, based on experience for example, as probabilities in the reasoning. However, if data are available to inform these beliefs it is sensible to take advantage of these data. For this example, the following data are available.

Worldwide, at least one in three people in most countries are diagnosed with a mental disorder at some point in their life (Andrade et al., 2000). Prevalence of psychiatric disorders varies considerably across countries. Thus, in what follows, figures more appropriate to a particular country should be used if that is felt necessary. The figures used here are primarily for illustration. Thus, the probability for a person of a diagnosis of a psychiatric disorder at least once in a lifetime is taken to be 0.33 (as an approximation to one-third). Formulaically this can be denoted as $Pr(A^+) = 0.33$ where Pr denotes `Probability'. This is the uncertainty associated with the information of relevance to the event *S*, whether a person will (S^+) or will not (S^-) commit suicide.

Suicide (S^+) has an average population rate of 1.2 per 10000 in the United Kingdom (World Health Organization, 2014). This can be written, approximately, as $\Pr(S^+) = 0.0001$, hence approximating 1.2 in 10,000 with 1 in 10,000 for ease of explanation. Uncertainty about whether a person will commit suicide or not is represented by a rate from a survey. In the absence of any appropriate survey, the uncertainty can be represented by a measure of belief expressed as a probability.

The interpretation of $\Pr(S^+) = 0.0001$ is that, other things such as gender and environment, being equal, 1 in 10,000 people commit suicide and this is also interpreted as a comment about a particular person that the `probability of a person committing suicide (S^+) is 0.0001 or 1 / 10,000'.

Further studies in high-income countries have consistently found that at least 90% of those who commit suicide, have been diagnosed with a psychiatric disorder at least once in their lifetime (Hawton and van Heeringen, 2009). This is a conditional probability of an event in light of some evidence (S^+ that one has committed suicide) and may be denoted as $\Pr(A^+|S^+) = 0.9$. This notation is read as `the probability a person has been diagnosed with a psychiatric disorder at least once in their lifetime (A^+) given that (the vertical bar |) they have committed suicide (S^+) is 0.9 or 9/10 or 90%. It is important to note that whilst 90% of those who committed suicide had been diagnosed with a psychiatric disorder at least once in their lifetime it is not the case that 90% of those who have been diagnosed with a psychiatric disorder at least once in their lifetime will commit suicide. However, the probability of interest for the diagnostician is not the probability that person who has committed suicide has been diagnosed with a psychiatric disorder. It is the transpose of this, namely the probability a person will commit suicide if they have been diagnosed with a psychiatric disorder at least once in their lifetime. Given the probabilities above, this probability of interest can be determined. The information available is listed below.

- S^+: a person will commit suicide;
- S^-: a person will not commit suicide;
- A^+: a person has been diagnosed with a psychiatric disorder at least once in their lifetime
- $\Pr(A^+) = 0.33$;
- $\Pr(S^+) = 0.0001$;
- $\Pr(S^-) = 0.9999$. The two events `will commit suicide' and `will not commit suicide' are what is known as `mutually exclusive and exhaustive events'. They are mutually exclusive since one cannot both commit suicide and not commit suicide. They are exhaustive as one cannot partially commit suicide, either one does or one does not. Thus, one or other of S^+ and S^- is certain to occur and hence $\Pr(S^+ \; or \; S^-) = \Pr(S^+) + \Pr(S^-) = 1$. The corresponding general result is that the sum of the probabilities of mutually exclusive and exhaustive events is 1.

- $\Pr(A^+|S^+) = 0.9$.

The probability of interest is the probability that a person who has been diagnosed with a psychiatric disorder will commit suicide. This is $\Pr(S^+|A^+)$. It is known that 33% (or probabilistically $\Pr(A^+)$) of the population have a psychiatric disorder. The probability of interest is the proportion of those that will commit suicide. It is known that 0.01% ($\Pr(S^+)$) of the population will commit suicide and of that 0.01%, 90% ($\Pr(A^+|S^+)$) will have been diagnosed with a psychiatric disorder. Thus 90% of 0.01% have both committed suicide and been diagnosed with a psychiatric disorder, denoted $\Pr(A^+ \; and \; S^+)$. We are interested in the proportion of the population with diagnosis of a psychiatric disorder that will commit suicide. The probability of interest can be calculated using Bayes theorem[1] , with C in the footnote replaced by S+ and D in the footnote replaced by A+',as

$$\Pr(S^+|A^+) = \frac{\Pr(A^+ \; and \; S^+)}{\Pr(A^+)} = \frac{\Pr(A^+|S^+) \times \Pr(S^+)}{\Pr(A^+)} = \frac{0.99 \times 0.0001}{0.33} \cong 0.00027$$

The general result concerning mutually exclusive and exhaustive events applies to events conditioned on other events. In this example the conditioning event is A and $\Pr(S^+|A) + \Pr(S^-|A^+) = 1$ and so $\Pr(S^-|A) = 1 - \Pr(S^+|A^+) = 0.99973$. Figure 1 shows a probability tree diagram for this example.

The probability 0.00027 (or 0.027%) should be compared with that of 90% for the proportion of those who committed suicide that had been diagnosed with a psychiatric disorder at least once in their lifetime. The error of confusing $\Pr(A^+|S^+)$ with $\Pr(S^+|A^+)$ is known as the *prosecutor's fallacy* (Thompson and Schumann, 1987).

The result

[1] Bayes theorem describes the probability of an event, e.g. C, based on prior knowledge, e.g. event D, of conditions that might be related to the event C. This is represented formulaically as $Pr(C|D) = \frac{\Pr(D|C) \times P(C)}{P(D)}$

$$\Pr(S^+|A^+) = \frac{\Pr(A^+|S^+)\ \times \Pr(S^+)}{\Pr(A^+)}$$

applies to S^- also.

$$\Pr(S^-|A^+) = \frac{\Pr(A^+|S^-)\ \times \Pr(S^-)}{\Pr(A^+)}$$

Division of the result for S^+ by the result for S^- gives a result known as the odds form of Bayes' theorem.

$$\frac{\Pr(S^+|A^+)}{\Pr(S^-|A^+)} = \frac{\Pr(A^+|S^+)}{\Pr(A^+|S^-)} \times \frac{\Pr(S^+)}{\Pr(S^-)} \qquad (1)$$

, where the common term $\Pr(A^+)$ cancels out.

This result demonstrates the updating of the prior odds $Pr(S^+)/Pr(S^-)$ in favour of a prognosis S^+ (odds prior to information related to the event) to posterior odds $\Pr(S^+|A^+)$ / $\Pr(S^-|A^+)$ in favour of the prognosis in the light of information *A*. The ratio $\Pr(A^+|S^+)\ /\Pr(A^+|S^-)$ is known as the *likelihood ratio (LR)*. It is the factor which updates prior odds $\Pr(S^+)\ /\Pr(S^-)$ in favour of an event to posterior odds $\Pr(S^+|A^+)$ / $\Pr(S^-|A^+)$ in favour of an event in the light of new information and it may be thought of as the *value* of the information. A LR greater than 1 supports one proposition and a LR less than 1 supports the other. The closer LR is to 1 (either less than or greater than 1) the weaker the support for the event e.g. for S^+ (LR > 1) or for S^- (LR < 1)

Equation (1) can also be expressed as

$$posterior\ odds = likelihood\ ratio \times prior\ odds.$$

It is possible with a rearrangement of (1) to determine $\Pr(A^+|S^-)$ and hence the LR.

$$\Pr(A^+|S^-) = \frac{\Pr(A^+|S^+)\times\Pr(S^+)\times\Pr(S^-|A^+)}{\Pr(S^+|A^+)\times\Pr(S^-)} = \frac{0.9\times0.0001\times0.99973}{0.00027\times0.9999} = 0.3332767.$$

$$LR = \frac{\Pr(A^+|S^+)}{\Pr(A^+|S^-)} = \frac{0.9}{0.3332767} = 2.7.$$

Verbally this result maybe expressed as follows. The odds in favour of committing suicide for a member of the general population (prior odds) are 1 in 9999. More specifically, $\frac{\Pr(S^+)}{\Pr(S^-)} = \frac{1/1000}{999/1000}$. If such a person has a diagnosis of a psychiatric disorder, the odds in favour of committing suicide (posterior odds) increase by a factor of 2.7.

One may argue that the posterior odds are increased by a factor of 170% (from 1 in 9999 to 2.7 in 9999). This result may appear impressive but the increase is from a very small number to another very small number. Relative measures are very informative in comparing an event across two groups (diagnosed vs non-diagnosed) but they do not reveal any information about the likelihood or the odds of the event in the two groups. Suicides are rare and the probability of suicide is still very small regardless of the diagnosis of a psychiatric disorder. Given an LR of 2.7 and in order for the posterior odds to exceed one (so that it is more likely for a person to commit suicide, than not commit suicide given a diagnosis with a mental disorder), prior odds should be at least 0.37 (because $0.37 \times 2.7 = 1$) which equates to an absolute risk of committing suicide at least equal to 0.27 ($P(S^{+}) \geq 0.27$), which is certainly an unrealistic probability for a person in the general population about whom nothing is known to commit suicide. All the probabilistic measures entertained in this example are presented in Table 1.

Example 2: Relationship of a symptom to a diagnosis

It is also possible using the same arguments and relationships such as the odds version of Bayes' theorem to update the probability a person has schizophrenia (a diagnosis) given evidence of a symptom. The example presented here uses diagnostic test accuracy studies (Takwoingi, Riley and Deeks, 2015) and a blood-based laboratory test for the diagnosis of schizophrenia (Schwarz et al., 2010). This test has a sensitivity (proportion of confirmed schizophrenia cases correctly identified by the test) and specificity (proportion of confirmed non-schizophrenia cases correctly identified by the test) that are both equal to 0.83. For comparison with the ideas given in Example 1:

- *Diagnosis*: The diagnosis is that a person has schizophrenia. This is the event about whose occurrence or not the clinician is uncertain. Denote this event as S^{+}. The event that a person will not have schizophrenia is denoted S^{-}.
- *Symptom*: This is the information of relevance to the event. The symptom is the result of the laboratory test. In contrast to example 1, a distinction is drawn between a positive result, denoted A^{+}, and a negative result, denoted A^{-}.

The information available is summarised below.

- S^{+}: a person has schizophrenia;
- S^{-}: a person does not have schizophrenia;
- A^{+}: the laboratory test has a positive result;
- A^{-}: the laboratory test has a negative result.
- $Pr(A^{+}|S^{+})$: sensitivity, the probability a person with schizophrenia has a positive test result, here equal to 0.83;
- $\Pr(A^{-}|S^{-})$: specificity, the probability a person without schizophrenia has a negative test result, here equal to 0.83;
- $\Pr(S^{+}) = 0.005$ the proportion of the population diagnosed with schizophrenia (World Health Organization, 2014)
- $\Pr(S^{-}) = 0.995$: The two events `has schizophrenia' and `does not have schizophrenia' are mutually exclusive and exhaustive events. Thus $\Pr(S^{-}) = 1 - \Pr(S^{+})$.

The probability a non-psychotic patient (one without schizophrenia) has a negative test result is 0.83. Thus, the probability this non-psychotic patient has a positive test result is $\Pr(A^{+}|S^{-}) = 0.17$. (Tests are assumed to have only two possible outcomes, positive and negative, which are mutually exclusive and exhaustive. No results are inconclusive.) Similarly, the probability a psychotic patient has a negative test, $\Pr(A^{-}|S^{+}) = 0.17$.

The overall proportion of the population that respond positively to the test is 83% of the 0.5% that have schizophrenia and 17% of the 99.5% that do not. This proportion is 17.33% of the population $((0.83 \times 0.005 + 0.17 \times 0.995) \times 100)$. Of that 17.33%, the number that respond positively, the number that have both schizophrenia and a positive test result, is 0.5% of 83% or 0.415%, a proportion 0.415/17.33 (0.024 or 2.4%) of the population that respond positively to the test. Similarly, of that 17.33%, the number that do not have schizophrenia but do have a positive test result, is 99.5% of 17% or 16.915%, a proportion 16.915/17.33 or 97.6% of the population that respond positively to the test and do not have schizophrenia. The initial proportion of the population with schizophrenia was 0.5%. The additional information of a positive test result has raised this proportion to 2.4% or an increase by a factor of 4.8. The factor can be shown to be the ratio of sensitivity to (1 - specificity) (0.83/0.17 = 4.79) and is known as a positive likelihood ratio in diagnostic test studies[2]. It represents the increase in odds favouring the outcome (e.g. schizophrenia) given a positive test result. A positive test result is 4.79 times as likely for a psychotic patient as for a non-psychotic one. LRs greater than one indicate that the test result is associated with the condition. This result can be shown mathematically using the odds form of Bayes theorem (1) where the notation refers to Example 2. Prior and posterior odds are also called pre-test and post-test odds in diagnostic studies. Hence

$$\frac{\Pr(S^+|A)}{\Pr(S^-|A)} = \frac{\Pr(A|S^+)}{\Pr(A|S^-)} \times \frac{\Pr(S^+)}{\Pr(S^-)} = \frac{0.83}{0.17} \times \frac{0.005}{0.995} = 4.79 \times 0.005 = 0.00415/0.16195$$
$$= 0.024$$

Addition of an additional and independent test

Consider a second test for schizophrenia independent of the second test, again with sensitivity and specificity 0.83, independent of the first test. Denote a positive result of this test as B^+ and a negative result as B^-. As with test A, $\Pr(B^+|S^+) = \Pr(B^-|S^-) = 0.83$ and $\Pr(B^+|S^-) = \Pr(B - |S+= 0.17)$. The proportion of people with positive test results for both A and for B is not the product of the proportion of people with a positive test result for A and the proportion of people with a positive test result for B.

A person gives a positive result for both tests. The probability of interest is still that the person has schizophrenia. It has been shown that for a positive result for test A, the probability is 0.024. For two positive results from independent tests with the same specificities and sensitivities of 0.83, the probability a person has schizophrenia is not 0.024^2 or 0.0006. Mathematically, $\Pr(S^+|A^+, B^+) \neq \Pr(S^+|A^+) \times \Pr(S^+|B^+)$. Such a result would suggest additional positive results led to a decrease in the probability of schizophrenia. An explanation as to why this is not so is given in the following paragraph.

For those with a positive result to the first test, only 2.4% of the population have schizophrenia. Thus, the overall proportion of the population that respond positively to the second test as well as positively to the first test is 83% of the 2.4% that have schizophrenia and 17% of the 97.6% that do not. This is 18.6% of the population ($(0.83 \times 0.024 + 0.17 \times 0.976) \times 100$). Of that 18.6%, the number that have both schizophrenia and a positive test result to both tests is 2.45% of 83% which equals approximately 10% (0.83 x 0.024/0.186 = 0.107. The initial proportion of the population with schizophrenia was 0.5%. The additional information of a first positive test result has raised this proportion to 2.4% or an increase by a factor of 4.9. A positive result to a second test, independent of the first raises the proportion to 10%. The probability a person with positive result to both tests has schizophrenia is 10%, compared with the base rate 0.5%, an increase by a factor of 20.

Comparison with the evaluation of evidence in criminal justice

Similar arguments to those in these two examples occur in the evaluation and interpretation of evidence in forensic science in the administration of criminal justice. Instead of presence or absence of a disease (e.g. schizophrenia) or the fulfilment or not of a prognosis (suicide), there

[2] All likelihood ratios are positive. In the context of our example, a diagnostic test accuracy study, a positive likelihood ratio is a standard terminology that refers to the increase in odds favouring the outcome given a positive test result

are two propositions, put simply as the defendant is guilty or the defendant is innocent. In this context, these propositions are those of true guilt and true innocence, not the verdicts `Guilty' or `Not guilty' of a jury. Instead of the test result (positive or negative) or the diagnosis (psychiatric disorder) there is scientific evidence. Consider evidence that the DNA profile of a defendant matches in some sense that of a blood stain found at a crime scene. If the defendant were innocent, this match would be fortuitous and the probability of a match by chance would be very small. The probability of finding the evidence of a match if the defendant is innocent is very small. However, this probability is not to be equated to the probability that the defendant is innocent if there is a match of DNA profiles. More generally, the probability of innocence for someone on whom incriminating evidence has been found is not the same as the probability of incriminating evidence being found on a person who is innocent. The claim that these two probabilities are equal is another example of the prosecutor's fallacy (Thompson and Schumann, 1987). Similarly, a large probability of a positive test result for someone who has a disease should not be equated to a large probability that a person with a positive test result has the disease.

Conclusion

Probabilistic reasoning is not innate and relies on good education. The effects of a diagnosis on a prognosis and of a symptom on a diagnosis are discussed using an approach based on an LR. The LR provides a continuous measure of support for one proposition (prognosis or diagnosis) based on an event (diagnosis or symptom). It is also shown how to combine the results from two independent events. Notice that the LR is said to provide support for one proposition over another. The LR does not comment on the truth or otherwise of a particular proposition. In order to do that, knowledge of the prior odds $\frac{\Pr(S^+)}{\Pr(S^-)}$ is needed. In a similar spirit, lots of medical journals, including BMJ, require that along with relative measures (e.g. relative risks), absolute measures of the event in each group (absolute risks) should be reported.

References

1. Andrade L, Caraveo-Anduaga JJ, Berglund P, *et al*. Cross-national comparisons of the prevalences and correlates of mental disorders. *Bull World Health Org* 2000;**78**:413-26. doi.org/10.1590/S0042-96862000000400003
2. Hawton K, van Heeringen K. Suicide. *The Lancet* 2009;**373**:1372-81. doi.org/10.1016/S0140-6736(09)60372-X.
3. Lindley DV. Understanding uncertainty, Revised edition. John Wiley and Sons Inc, Hoboken, New Jersey; U.S.A 2014
4. Schwarz E, Izmailov R, Spain M, *et al*. Validation of a blood-based laboratory test to aid in the confirmation of a diagnosis of schizophrenia. *Biomark Insights* 2010;5:39-47. doi.org/10.4137/BMI.S4877
5. Takwoingi Y, Riley RD, Deeks JJ. Meta-analysis of diagnostic accuracy studies in mental health. *Evid Based Ment Health* 2015;**18**:103-9. doi.org/10.1136/eb-2015-102228.
6. Thompson WC. Schumann EL. Interpretation of statistical evidence in criminal trials. The prosecutor's fallacy and the defence attorney's fallacy. *Law Hum Behav* 1987;11:167-87.
7. World Health Organization (2014). WHO | Schizophrenia. Schizophrenia Fact Sheet. Retrieved from http://www.who.int/mental/health/management/schizophrenia/en/.
8. World Health Organization (2014). Preventing suicide: a global imperative. World Health Organization. http://www.who.int/iris/handle/10665/131056

Table 1: Interpretation and formulae for various probabilistic measures. We use upper indices + and – to denote whether event A happens or not. For event B we assume that is always happen and we omit the upper index.

Measure	Interpretation	Formula
Probability (e.g. of an event A^+).	A measure of the likelihood of an event. It takes a number between 0 (impossible) and 1(certain)	$\Pr(A^+)$
Probability of a complementary event (e.g. A^-)	A measure of the likelihood that an event will not occur	$\Pr(A^-) = 1 - \Pr(A^+)$
Conditional probability (e.g. of an event A given an event B)	A measure of the likelihood of an event (e.g. A) given that another event (e.g. B) has occurred	$\Pr(A^+\|B)$
Bayes theorem	Bayes theorem describes the probability of an event, e.g. A^+, based on prior knowledge, e.g. event B, of conditions that might be related to the event A^+.	$\Pr(A^+\|B) = \frac{\Pr(B\|A^+) \times \Pr(A^+)}{\Pr(B)}$
Prior Odds (e.g. of an event A)	The odds in favour of A; the probability A will occur divided by the probability it will not occur.	$prior\ odds = \frac{\Pr(A^+)}{\Pr(A^-)} = \frac{\Pr(A^+)}{1 - \Pr(A^+)}$
Posterior Odds (e.g. of an event A given that B has occurred)	The odds of A in light of B; the probability A will occur given B has occurred divided by the probability A will not occur given that B has occurred. They inform us how odds of A have been updated given that B has occurred.	$posterior\ odds = \frac{\Pr(A^+\|B)}{\Pr(A^-\|B)} = \frac{\Pr(A^+\|B)}{1 - \Pr(A^+\|B)}$
Likelihood Ratio	The factor which updates prior odds in favour of an event (e.g. A+) to posterior odds in favour of an event in the light of new information (e.g. B).	$LR = \frac{P(B\|A+)}{P(B\|A^-)}$