The nine circles of EBM hell

Evidence-based medicine (EBM), after being defined as “the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients”, was subsequently redefined as “the integration of best available evidence, clinical expertise and patient preferences and values” (Sackett et al., 2000).
Evidence means scientific proof, based on data collection and statistical analysis. From this perspective, the field of neurology is a particularly tricky area in which there is an ever-present risk of over-emphasising the importance, especially in a translational sense, of the results of neuroscience studies, even when these results are actually still a long way from representing a solid scientific basis on which to bring about real improvements in neurorehabilitation and other types of neuro-intervention, and in the quality of life of neurological patients.
In this editorial, we take a humorous look at the errors neuroscientists need to avoid, drawing an analogy with the nine sins described by Dante in his *Inferno*. A similar approach has already been adopted in a previous paper published in *Perspectives on Psychological Science* (Neuroskeptic, 2012), but whereas that paper focused mainly on the issue of immoral, arbitrary use of data, we instead focus on the most common errors that neuroscientists make.

Circle 1: Limbo – the circle of narrative scientists
The first circle of Dante’s hell is the place of virtuous non-Christians and unbaptised pagans, and it is peopled by important personages such as Homer, Socrates, Aristotle, Cicero, and Hippocrates. Our “hell”, on the other hand, is inhabited by very important neuroscientists who were sceptical about EBM, or preferred narrative medicine, such as Oliver Sacks, author of some amazing books. The field of neuroscience itself might actually be said to have been born from early narrative studies of particular cases, such as that of Phineas Gage, and in more recent times narrative medicine has found itself back in vogue, attracting serious attention among health care professionals (Morris et al., 2008). In their famous systematic review highlighting the lack of randomised controlled trials dealing with the use of parachutes to prevent death related to gravitational challenge, Smith and Pell (2003) pointed out that absence of evidence is not related to absence of efficacy. They are, of course, right in saying this, but there is no doubt that, for patients, the presence of evidence is, at once, a demonstration and a guarantee of efficacy. In short, narrative medicine is very appealing and intriguing, and it is true that not all the effective treatments are proved by EBM; however, it is also true all treatments based on EBM are effective, until proven otherwise (and providing the following errors were avoided).

Circle 2: Lust – the circle of the overly passionate who overstate their results
Narrative medicine can certainly teach EBM something about the passionate reporting of results, although it is important to remember that passionate reporting cannot overcome the limits of hard data. Too often, neuroscientists speculate about their results, and while this sometimes helps to drive science forward, at other times, it leads to the reaching of conclusions that are unsupported by results. Findings that are reported using speculative language, or in exaggerated terms, need further evidence before they can be applied, in clinical practice, to patients. The increasing pressure on scientists to publish, obtain grants and write higher profile papers can lead them to exaggerate the importance of their own work, and this behaviour, reflected in sensationalised headlines and findings or catchy titles (such as that of this editorial), may generate a bias (Fanelli, 2010). Passion is fundamental in scientists, but science should provide evidence, not just hope.

Circle 3: Gluttony – the circle of those who overindulge in measures
Scientists who measure hundreds of clinical scores and parameters, and then try every statistical test until they get a p-value less than 5%, are forgetting Bonferroni’s rules and also Galileo’s scientific method (Neuroskeptic, 2012). The guidelines of many scientific journals now require Authors to identify just one main objective of their study, and just one relevant primary outcome measure to be used to evaluate its main hypothesis.

Circle 4: Greed – the circle of the prodigal, that is, researchers who enrol too many subjects (and the similar problem of enrolling too few)
Friston (2012), in his ten ironic rules for non-statistical reviewers, suggested that manuscripts should be rejected if the sample concerned is too small, but also when it is too large. In the former case, any significant results cannot be considered sufficiently solid, while non-significant results could lead to a type II error, wherein a null hypothesis is confirmed, even though in reality it is false. However, an excessively large sample is also to be avoided: it increases the possibility of a type I error and of obtaining a p-value lower than 0.05 even for a minimal, clinically non-significant change, which would have the effect of exposing too many patients to a treatment whose efficacy is still to be proven. There exist specific criteria for calculating sample size, and this process should be taken into account right from the planning stage of the study.

Circle 5: Anger – the circle of those who, maddeningly, employ scales that lack sensitivity or other psychometric features
Sensitivity is an important trait in human beings, and it is also an important feature in clinical scales and measures. Scientists should use only clinical scales and measures whose psychometric features have been clearly proven. There are many psychometric features to take into account, such as validity, reliability, accuracy, precision, specificity, sensitivity, and so on.

Circle 6: Heresy – the circle of scientists who advance implausible biological or physiological theories
To prove an association or a causation it is not sufficient to find a coefficient of correlation with a p-value lower than 0.05. Hill's criteria highlighted the importance of biological plausibility and coherence (Hill, 1965), while Evans (1976), revisiting Koch's postulates, highlighted that every finding should have a biological and epidemiological meaning. Matthews (2000), in the paper "Storks deliver babies (p=0.008)", provided a hilarious illustration of the point that correlation *per sé* is not sufficient.

Circle 7: Violence – the circle of those who mistreat numbers and data
Pythagoras said that numbers rule the universe. At the same time, numbers have their rules. Most scientists are reluctant to use non-parametric statistics, even when the distribution of their data is not normal. Three out of four types of variables (nominal, ordinal, discrete) are often not normally distributed and therefore need non-parametric tests (only continuous variables are often normally distributed, but not always). Another "violence" is the selective reporting of data, classifying as outliers those not agreeing with the theory being tested.

Circle 8: Fraud – the circle of deceptive scientists
Without necessarily behaving in a fraudulent manner, scientists sometimes perform studies without completely declaring all the possible ethical issues. Such issues might include: possible conflicts of interests, the receipt of funding (in the form of money or devices) from companies, overstatement of findings, allocation of subjects in a manner that is not truly randomised, use of non-blind assessors in studies claiming to be blind, randomised controlled trials, and partial description of methods (making the study impossible to replicate).

Circle 9: Treachery – the circle of the untrustworthy
Scientists require that people trust in science, but to earn that trust, science should be transparent, replicable, verifiable and ethical. In the past, many studies were not ethical. All neurological studies in human subjects should be performed in accordance with the principles of the Declaration of Helsinki and approved by an ethics committee.

In conclusion, the errors that neuroscientists can run into should not be seen as a proof of the fallacy of EBM. Instead, it should be appreciated that they highlight areas where EBM can be improved. In this editorial we have shown that EBM has something to learn, in terms of form, from narrative medicine, but it absolutely must maintain, in its content, the rigor that science demands.

Marco Iosa[a], Gabriella Antonucci [a,b], Stefano Paolucci[a]
[a]*IRCCS Fondazione Santa Lucia, Rome, Italy*
[b]*Department of Psychology, Sapienza University of Rome, Rome, Italy*
E-mail: m.iosa@hsantalucia.it

References

Evans AS (1976). Causation and disease: the Henle-Koch postulates revisited. Yale J Biol Med 49: 175-195.
Fanelli D (2010). Do pressures to publish increase scientists' bias? An empirical support from US States Data. PLoS One; 5:e10271.
Friston K (2012). Ten ironic rules for non-statistical reviewers. Neuroimage 61: 1300-1310.
Hill AB (1965). The environment and disease: association or causation? Proc R Soc Med 58: 295-300.
Matthew R (2000). Storks deliver babies (p= 0.008). Teaching Statistics 22: 36-38.
Morris DB (2008). Narrative medicines: challenge and resistance. Perm J 12: 88-96.
Neuroskeptic (2012). The nine circles of scientific hell. Perspect Psychol Sci 7: 643-644.
Sackett DL, Straus SE, Richardson WS, et al (2000). Evidence-Based Medicine: How to Practice and Teach EBM (2nd edition). Churchill Livingstone, New York.
Smith GC, Pell JP (2003). Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. BMJ 327:1459-1461.