BIOENGINEERING SEMINAR

“The Assessment of Science: the Relative Merits of Post-Publication Review, the Impact Factor and the Number of Citations”

Tuesday – March 11, 2014 – 10:15 a.m.
EPFL – room SV1717a

Prof. Adam Eyre-Walker
School of Life Sciences, University of Sussex, Brighton (UK)

host: Prof. J. Jensen

Abstract

The assessment of scientific publications is an integral part of the scientific process. I will present analyses of three methods of assessing the merit of a scientific paper: subjective post-publication peer review, the number of citations gained by a paper, and the impact factor of the journal in which the article was published. I will show that there are moderate, but statistically significant, correlations between assessor scores, when two assessors have rated the same paper, and between assessor score and the number of citations a paper accrues. However, assessor scores are more strongly correlated to the impact factor of the journal in which the paper is published than they are to each other or to the number of citations. This might be because assessors are overrating papers in high impact journals. If we control for this potential bias, we find that the correlation between assessor scores and between assessor score and the number of citations is weak, suggesting that scientists have little ability to judge either the intrinsic merit of a paper or its likely impact. I also show that the number of citations a paper receives is an extremely error-prone measure of scientific merit. Finally, I will argue that the impact factor is likely to be a poor measure of merit, since it depends on subjective assessment. I conclude that the three measures of scientific merit considered here are poor; in particular subjective assessments are an error-prone, biased, and expensive method by which to assess merit. I argue that the impact factor may be the most satisfactory of the methods we have considered, since it is a form of pre-publication review. However, I will emphasise that it is likely to be a very error-prone measure of merit that is qualitative, not quantitative.