Open Access, Impact, and Demand: Why Some Authors Self-Archive Their Articles

4 Barker DJP, Gluckman PD, Godfrey KM, Harding JE, Owen JA, Robinson JS. Fetal nutrition and cardiovascular disease in adult life. *Lancet* 1993; 341:938-41.
5 Frankl S, Elwood P, Sweetnam P, Yarnell J, Davey Smith G. Birthweight, body mass index in middle age and incidence of coronary heart disease. *Lancet* 1996;348:1478-80.
6 Eriksson JG, Forsén T, Tuomilehto J, Winter PD, Osmond C, Barker DJP. Catch-up growth in childhood and death from coronary heart disease: longitudinal study. *BMJ* 1999;318;427-31.
7 Rich-Edwards JW, Kleinman K, Michels KB, Stampfer MJ, Manson JE, Rexrode KM, et al. Longitudinal study of birth weight and adult body mass index in predicting risk of coronary heart disease and stroke in women. *BMJ* 2004;330:1115-8.
8 Eriksson JG, Forsén T, Tuomilehto J, Osmond C, Barker DJP. Early growth and coronary heart disease in later life: longitudinal study. *BMJ* 2001; 322:949-53.
9 Barghava SK, Sachdev HS, Fall CHD, Osmond C, Lakshmy R, Barker DJP, et al. Relation of serial changes in childhood body-mass index to impaired glucose tolerance in young adulthood. *N Engl J Med* 2004; 350:865-75.
10 Eriksson JG, Forsén T, Osmond C, Barker DJP. Pathways of infant and childhood growth that lead to type 2 diabetes. *Diabetes Care* 2003;26: 3006-10.
11 Eriksson J, Ylihärsilä H, Forsén T, Osmond C, Barker DJP. Exercise protects against Type 2 diabetes in persons with a small body size at birth. *Prev Med* 2004;39:164-7.
12 Eriksson JG, Lindi V, Uusitupa M, Forsén TJ, Laakso M, Osmond C, et al. The effects of the Pro12Ala polymorphism of the peroxisome proliferator-activated receptor-gamma2 gene on insulin sensitivity and insulin metabolism interact with size at birth. *Diabetes* 2002;51: 2321-4.

Open access, impact, and demand

Why some authors self archive their articles

Learning in practice p 1128

The great current divide in scientific publishing is between open access articles—that is, those freely available on the internet—and non-open access ones, those for which a reader has to pay on order to gain access to them. Before Jonathan Wren's study appeared (p 1128)[1] we knew that open access copies of scientific journal articles published in non-open access (subscription based) journals were a fairly small subset of the overall journal literature.[2] Wren studied just which subset it was and found that papers from journals with high impact factors were more likely to have free online copies at other locations around the web than papers from low impact journals.

To show why this matters, and why it's puzzling, let's review what we knew before Wren did his study. We knew that some scientists deposited copies of their published articles in open access repositories, a process called self archiving. We knew that about 80% of subscription based journals allowed their authors to do so.[3] Hence, we knew that self archiving was compatible with copyright and with publication in a non-open access journal. We knew that it took an author about 10 minutes to self archive one paper.[4] We knew that the open access archives where authors deposited articles were "interoperable," which means that they conformed to a common standard allowing users to search them all at once, as if they comprised one grand, virtual archive. We knew that there were many effective cross archive search tools to take advantage of this interoperability. We also knew that Google, Yahoo, and other mainstream search engines were indexing these archives. We knew that there were more than 400 standard compliant archives around the world,[1] with new ones launched every week. We knew that, because of their wider reach and increased visibility, open access articles were cited 50-300% more often than non-open access articles from the same journal and year,[5] although we still don't know how many authors and journals realise this. We knew, in other words, that self archiving was a small investment for authors with a large pay-off.

We knew that the practice of self archiving was catching on. But we also knew that proponents of open access were frustrated with the slow rate of its growth.[6] We knew that most publishing scientists were not opposed to open access but didn't know much about it or its benefits.[7] We knew that open access proponents wanted more authors to understand that self archiving was quick, easy, lawful, and beneficial.[8] Meantime, authors who did practise self archiving were steadily creating a critical mass of peer reviewed, open access research literature.

Wren's result matters because it gives us some insight into the motivation of authors who self archive. Authors with articles in high impact journals already have comparatively large audiences. They might be seeking even larger audiences (open access articles reach a much larger set of readers than any priced journal, in print or online). They might be showing off, posting copies to display their success in having been accepted by a prestigious journal. They might be practising what media scholars call "push," bringing their work to the attention of those who might not know about it, even though those recipients already had free online access to it. These are all different ways of saying that self archiving authors were advertising themselves and their work. This is not a cynical diagnosis. On the contrary, this kind of notice can advance research in the author's niche and advance the author's career.

It's possible that many of these free online copies were posted by readers, not authors, though Wren has no data on this. For convenience, I'll assume that reader posting was the exception rather than the rule, but this might oversimplify the analysis. What's puzzling is that authors who publish in low impact journals turn to open access at lower rates. It seems that they have the same interest in enlarging their audience and impact as authors who publish in high impact journals, if not more. One possibility is that they are not proud of where they published and fear that the "advertisement" would be double edged.

Another possibility is that more high impact journals than low impact journals give authors permission for self archiving. Wren didn't investigate this possibility, but he did name the 13 (non-open access) journals he chose to study. I looked up their self archiving policies and found that the high impact journals on his list were indeed more likely to permit self archiving than the low impact journals. However, most of the high impact journals did not permit archiving of the published PDF, and Wren studied only free online PDFs. Hence, this alluring alternative explanation largely disappears, and we're back to the puzzle.

Wren made another, even more enigmatic discovery. Articles from open access journals were just about

BMJ 2005;330:1097–8

as likely to have free online copies elsewhere online as articles from non-open access journals. What's puzzling is that authors would provide open access for articles that were already open access. One possibility is that this is still self advertising. Authors may put copies where they are more likely to be seen, even if existing copies sufficed for readers who ran searches or knew where to look. Another possibility is that the free online copies were posted by readers, not by authors. When I've found readers copying and reposting my own articles, some told me that they wanted more assured access, not knowing how long the originals would remain freely available.

Some journals deposit their own articles in open access repositories to assure their long term preservation and accessibility. But Wren's study included only one journal—the *BMJ*—with such a policy. Hence, author and reader deposits will still have to account for the bulk of the free online copies that Wren studied. Wren's data show a steady upward trajectory over the past decade for open access copies of journal articles retrievable by Google searches, his most encouraging result. This suggests that author self archiving is increasing, reader reposting is increasing, or "link rot" is making older copies less visible—most likely some of each.

One way that Wren summarises his conclusion needs some elaboration. He says, "Decentralised sharing of scientific reprints through the internet creates a degree of de facto open access that, though highly incomplete in its coverage, is none the less biased towards publications of higher popular demand." This is accurate but may leave the impression that most high demand articles are open access somewhere, when all we know so far is that most open access articles in the set he studied were high demand. It's possible that the vast majority of high demand articles are not yet open access, and indeed this seems likely. Most publishing scientists do not yet self archive their work and their reasons seem entirely unrelated to the demand, impact, or quality of their work—that is, they know too little about self archiving or believe they are too busy.

This is important because we ought to use Wren's results to understand why authors self archive and how to appeal to authors who don't. One lesson is that existing open access is demand driven to some degree. But this doesn't mean there is little or no unmet demand. On the contrary, unmet demand may be the norm, just as the sale of food is demand driven while the unmet demand exists in catastrophic proportions.

Peter Suber *research professor*

Earlham College, Richmond, Indiana 47374, USA
(peters@earlham.edu)

Competing interests: None declared.

1 Wren J. Open access and openly accessible: a study of scientific publications shared on the internet. *BMJ* 2005;330:1128-31.
2 University of Michigan Digital Library Production Service (OAIster). http://oaister.umdl.umich.edu/o/oaister/viewcolls.html (accessed 5 May 2005).
3 eprints.org sites at Southampton serving Open Archives. *Summary statistics so far*. http://romeo.eprints.org/stats.php (accessed 6 May 2005).
4 Carr L, Harnad S. *Keystroke economy: a study of the time and effort involved in self-archiving, a preprint*. Southampton, 2005. http://eprints.ecs.soton.ac.uk/10688/ (accessed 6 May).
5 Hitchcock S, on behalf of the Open Citation Project. *The effect of open access and downloads ('hits') on citation impact: a bibliography of studies*. Southampton, 2005. http://opcit.eprints.org/oacitation-biblio.html (accessed 6 May 2005).
6 Harnad S. Fast-forward on the green road to open access. *Ariadne*, January 2005. http://arxiv.org/ftp/cs/papers/0503/0503021.pdf (accessed 6 May 2005).
7 Rowlands I, Nicholas D, Huntingdon P. *Scholarly communication in the digital environment: what do authors want?* London: CIBER, 2004. http://opcit.eprints.org/oacitation-biblio.html (accessed 6 May.)
8 Swan A, Brown S. Authors and open access publishing. *Learned Publishing* 2004;17(3):219-24. http://cogprints.org/4123 (accessed 6 May 2005).

Monitoring surgical mortality

Scottish scheme has worked well but may not be transferable to other settings

Education and debate pp 1139, 1144

Should surgical mortality be routinely monitored? In this issue Thompson and Stonebridge present a compelling argument for systematic audits (p 1139)[1] and Esmail, in the first part of a new series on the General Medical Council and revalidation, argues that doctors will have nothing to fear from the GMC's revised plans (p 1144).[2]

The Scottish Audit of Surgical Mortality is a voluntary, peer reviewed, critical event analysis that has become an established part of standard surgical practice in Scotland. Scottish surgeons have shown tremendous support for the programme—99% of surgeons participate and 91% of deaths under surgical care in Scotland are audited. They support the scheme perhaps because it seems to be effective. After errors in specific processes of care (failure to use intensive care units and failure to use prophylaxis for deep venous thrombosis) were identified by the scheme as contributing to surgical deaths, system-wide changes occurred and the frequency of such errors greatly declined.

The potential effectiveness of a programme that focuses on death as the only critical event, however, may be limited. Although errors occur often in medicine,[3] errors contributing to death occur in only 6% of cases identified by Scottish Audit of Surgical Mortality. Errors that do not occur often or do not generally result in mortality are likely to be missed by such a programme. In addition, the focus of the programme on processes of care would indicate that feedback at the hospital level is at least as essential as feedback at the individual surgeon level.

The grassroots, clinician led model has worked well in Scotland but may not be easily transferable, particularly in settings where results of such a programme could have market influences. In the United States, most audits have taken the form of report cards where mortality (and in some cases morbidity) rates are calculated for a given procedure at the hospital or individual surgeon level, and the rates (generally adjusted for comorbidities) are compared between sites and surgeons. In such a system, attention is focused generally on the outliers who have poor results, with (in most cases) neither integrated analysis of the root cause nor any attempt to determine the processes of care that result in worse outcomes.

BMJ 2005;330:1098–9