

nature biotechnology

Numb and number

A good head for figures was important at last month's BIO 2004 in San Francisco, the annual bash of the US Biotechnology Industry Organization. It wasn't the profusion of 181 sessions, 900 speakers and hundreds of exhibitors that made the head spin, it was the observation that no one estimate of biotech sector performance is ever the same.

Let's try a simple figure. The number of companies involved in biotech in the United States is ...? Well it is 1,473 with 314 publicly held, according to accounting firm Ernst & Young; or it is 1,466 (318 public) according to BIO; or it is even 1,607 companies (393 public) according to our very own surveys (*Nature Biotechnology* directory 2003; *Nat. Biotechnol.* 22, 665–670, 2004). The US Department of Commerce generously estimates that there are more than 3,200 US companies using biotech. This is all very discombobulating if you are one of the 146,100 or 194,600 or 886,000 people employed by the US biotech industry (depending on whether you have been counted by E&Y, BIO or the Battelle Memorial Institute/Fleishman-Hillard, respectively).

What accounts for the discrepancies in these figures? First and foremost, the numbers depend on how biotech is defined. What is a biotech company? Is it a firm relying heavily on life science R&D for future business? Should it include life science reagent suppliers? Does it encompass large companies like Eli Lilly that produce biologics as well as small molecules? And are database providers or clinical research organizations biotechs? It was John Bu'lock, a biochemist at the University of Manchester, who in the early 1980s described the essence of what biotech does as "turning chickens* into gold." Statistics in the field, if mishandled, can go a long way to reversing that Bu'lockian process.

Which is why it is with some trepidation that *Nature Biotechnology* is launching in this issue a section that provides what we regard the most reliable key numerical indicators of the status of the biotech sector. On p. 798, we present some of the key economic indicators for the biotech sector in 2003. In the forthcoming months, we will be presenting facts and figures on R&D funding, intellectual property, drug market indicators, technology transfer and venture creation, agbiotech, regulatory agencies and biotech research. Every four months, we will present stats on the biotech sector's performance each financial quarter, including venture capital investment, initial and follow-on public offerings, partnership deals and M&A data, stock indices (past quarter compared with previous year) and a list of biotech companies acquired. (The data format will be standardized so readers know what to expect.)

Twenty years ago, there were hardly any stats available on the biotech sector. There were few 'industry watchers' or consultancies to track the sector (no bona fide biotech sector existed) and no investment houses that specialized in funding life science ventures to collate financial data. Today, in an interconnected electronic world where, in theory at least, all information on one computer is only one download or one e-mail away from all other computers, gathering reliable and useful data on commercial biotech ought to be easy. And, of course,

collecting data is easy. But gathering meaningful data and gathering it meaningfully is actually made much harder by its overwhelming supernumeracy.

It will come as absolutely no surprise to our readers that data, although important, are not as important as their provenance. In the average scientific report, data are only a fraction of the paper. The rest of the article qualifies that data, spelling out who are the collectors of the data, where they work, the design of the experiments that generated the data, the scientific and social context that makes the experiment worthwhile and the data's implications (and in the modern context, whether the results might have been influenced by authors' commercial associations).

When it comes to 'industry' statistics in biotech, all that care and attention is often discarded. Thus, although a comparison of the number of biotech companies in country X with country Y might seem easy, in all likelihood the counts in the different countries will not be comparable because they are drawn together by different sets of officials, possibly for very different reasons. In biotech, one government's statistics are not always comparable with another's. One hardly ever knows what the numbers represent or why they were gathered.

Even with a good or consistent definition, there is still the problem of keeping up with the flux in biotech. Numbers vary because the sector is always on the move. There are mergers and acquisitions and bankruptcies on the one hand; new ventures on the other. Estimating the number of startups is always tricky because many of these companies are so young and small they may not hit the statistician's radar for many years.

What applies to counts of company numbers also applies to counts of employees, R&D spending, investment and a panoply of other parameters. Even the major management accountancy firms are culpable of adding apples and pears. These are firms that would never dream of adding francs to euros to dollars in compiling the accounts of a multinational company. But in adding together national biotech profiles to get a global or European picture, this, in effect, is precisely what they do do.

And then there is the question of whether the data are worthwhile in any case. According to a recent analysis from biotech data specialists, Critical I, Iceland has only five companies employing about 600 people between them, whereas the United States has around 1,800 that employ about 200,000 people (yet another estimate). On the surface, the United States' sector completely dwarfs that of Iceland. But, because the population of Iceland is almost exactly 1,000th of that of the United States, the Atlantic island has three times the density of biotech employment. Indeed, on a pro rata basis, Iceland rivals the hottest of the US biotech hotspots, such as California and New England.

Readers should thus be wary when interpreting statistics on the biotech sector. As space permits, our charts will contain succinct definitions of the nature of the data and an indication of their source. And remember, these are just data: not everything that can be counted counts; and not everything that counts can be counted. 