Box S1 | Additional notes on data sources and analysis methodology

Note 1: identification of new drugs and sample definition. The Center for Drug Evaluation and Research (CDER) at the US FDA is responsible for approving all new molecular entities (NMEs, which are approved under New Drug Applications (NDAs)) and some new therapeutic biologics (NTBs, which are approved under Biologics License Applications (BLAs)). Until 2009, lists of new drugs approved each year by CDER were available at www.fda.gov/cder/rdmt/default.htm. These lists began in 1998 for NMEs and in 2004 for NTBs. NTBs approved between 1998 and 2003 were identified from a list at www.fda.gov/cder/biologics/biologics_table.htm. However, since 2009, the URLs have been changed. Lists of drugs approved year-by-year by the CDER can be generated from data files accessible at http://www.fda.gov/Drugs/InformationOnDrugs/ucm135821.htm. Lists for individual years are also available at URLs such as http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/DrugandBiologicApprovalReports/NMEDrugandNewBiologicApprovals/ucm081684.htm for 2001.

These URLs do not include biologics and vaccines approved by the Center for Biologics Evaluation and Research (CBER), which are listed at http://www.fda.gov/BiologicsBloodVaccines/ucm133705.htm. Regulation of NTBs was shared between CDER and the CBER in the period of this study. Prior to 2003, regulatory review of biologic products, including almost all monoclonal antibodies and protein therapies, was the responsibility of the CBER. In June 2003, regulatory authority over therapeutic biologics whose molecular composition is well defined (often referred to as biopharmaceuticals), including the vast majority of those made by recombinant technology, was transferred to CDER. CBER still regulates vaccines, allergenics (diagnostic kits and anti-sensitization substances), blood and blood products, blood test kits and related devices, and gene therapy, human tissue and xenotherapy products. Some of these clearly meet the definition of therapeutic biologics, in particular, recombinant blood products (such as clotting factors) and highly purified natural proteins whose main use is therapy for a particular disease; for example, alpha-1-proteinase inhibitor for patients with emphysema due to deficiency of this protein.

To make sample definition easier and more clear-cut, the main analysis considers only the CDER-regulated NTBs (including biopharmaceuticals approved pre-2003 by CBER). However, supplementary information S3 (box), note 19 summarizes the likely origins of nine recombinant or highly purified NTBs approved by CBER over the same ten-year period, and still regulated by CDER, whose development or manufacture embodies many of the advanced technologies used in CDER-regulated NTBs — four clotting factors, two purified alpha-1-proteinase inhibitors, two antidotes (to snake venom and digoxin) and one purified (non-recombinant) human protein C. The probable origins of these nine CBER-approved biologics were analysed to determine whether their inclusion in the main analysis would be likely to affect the study's main findings.

The sample definition method excludes compounds approved as new formulations of NMEs approved before 1998, even though the pre-1998 approval may have been for a completely different indication. So, Fujisawa's (now Astellas's) Protopic was excluded, even though it was the first non-steriodal topical medication for atopic dermatitis. The FDA approved Protopic in 2000 following standard review as a new formulation of tacrolimus, which the FDA first approved in 1994 following priority review as the oral medication Prograf to prevent transplant rejection. The selection method also excludes new esters, salts, or noncovalent derivatives of previously approved NMEs. So, Lexiva (fosamprenavir, approved 2003), a phosphate salt and prodrug of Agenerase (amprenavir, approved 1999) is excluded.

Note 2: expanded definition of new therapeutic biologics (NTBs). Most NMEs are small molecules with a molecular mass less than 1,000 daltons (1 kDa). The 37 NTBs approved under BLAs are polypeptide-based drugs, ranging in molecular mass from about 16 to 152 kDa. However, at least 21 of the drugs approved by the CDER as new molecular entities (NMEs) are also polypeptide-based drugs. With one exception — Mylotarg approved as an NME in 2000, which has a molecular mass of 152 kDa (but could be viewed as an NME

attached to a protein) — their masses range from 0.8 to 47 kDa, which is generally less than the drugs approved as NTBs, but more than almost all the other drugs approved as NMEs. The drugs approved by the CDER as NMEs also include an RNA aptamer, Macugen, which has a 50 kDa molecular mass, and one antisense oligonucleotide, Vitravene, whose 7.1 kDa molecular mass is probably heavier than any of the non-polypeptide/nucleotide NMEs.

Analysis of the 37 NTBs approved under BLAs showed that biotechnology companies and universities discovered nearly all of these drugs in the period studied. However, there likely are unique technical challenges in designing and developing protein and nucleotide based drugs[1-3], although the extent to which these challenges relate mainly to size, rather than biochemical structure, is not clear. To investigate trends for all polypeptide-based drugs, as well as polynucleotide-based drugs, a further analysis reclassified the 21 polypeptide-based NMEs, as well as the polynucleotide-based drugs Macugen and Vitravene as NTBs, to give a total of 60 NTBs under this expanded definition. The reclassified polypeptide drugs are indicated by NME* in the 4th column of supplementary information S2 (table), and the reclassified polynucleotide drugs are designated by NME** in the same column. Table 1 in the main text contains summary statistics for this expanded set of 60 NTBs.

Note 3: exclusion of selected agents approved by the FDA. Of the agents approved by the CDER in the period of this study, the following 22 NMEs (with approval dates in brackets) were excluded from the analysis:

• Eight diagnostic imaging agents: gadoversetamide (approved 1999), technetium TC 99M depreotide kit (1999), perflutren lipid microspheres (2001), dimyristoylphosphatidylcholine perflexane (2002), gadobenate dimeglumine (2004), trypan blue (2004), technetium 99m Tc fanolesomab (2006) and ammonia N13 (2007);

• Four variants of hyaluronidase to enhance absorption of intramuscularly administered drugs: Vitrase (2004), Amphase (2004), Hydase (2005) and Hylenex Recombinant (2005);

• Three compounds to enhance elimination of radioactive imaging agents or radioactive substances dispersed by a dirty bomb: Prussian blue (2003), pentetate calcium (2004) and pentetate zinc (2004);

• Two medicated contraceptive devices: NuvaRing and Ortho Evra (both 2001);

• An antidote against chemical warfare agents: perfluoroalkylpolyether and polytetrafluoroethylene (2000);

• A peritoneal dialysis solution: icodextrin (2002);

• A diagnostic agent that stimulates secretion of various digestive enzymes: human secretin, (2004);

• A common non-essential amino acid (L-glutamine, NutreStore) approved as an adjunct to recombinant growth hormone to stimulate the regrowth of intestinal mucosa in patients with short bowel syndrome (2006);

• A combination of previously known drugs to treat infection with *H. pylori*: Pylera (biskalcitrate; metronidazole and tetracycline hydrochloride; 2006).

Although Thyrogen (purified recombinant thyroid stimulating hormone, 1998) is approved only as an adjuvant to the diagnosis of persistent, metastatic thyroid cancer, it was retained because it has potential therapeutic applications and because its currently approved use permits continuation of normal thyroid replacement therapy in patients who have undergone thyroid resection/ablation because of cancer.

Thus the final set for analysis became 252 drugs, of which 215 are NMEs and 37 are NTBs.

Note 4: indicators of innovativeness. The FDA accords priority status to drugs that offer substantial benefit over currently marketed drugs. Priority status provides fast-tracking for regulatory approval, with the intention of completing review within six months. For drugs without priority status, the FDA aims to complete review within twelve months. So, priority status is an indicator of innovativeness with respect to medical needs. In this study, priority approved NMEs and standard approved NMEs are designated as pNMEs and sNMEs, respectively. Because FDA documents do not indicate which NTBs approved prior to 2004 received priority approval, this study usually analyses NTBs as a single category. Summary statistics for this classification are presented in Table 1.

Another indicator was created that emphasizes scientific innovativeness and risk-taking with respect to discovery. The criteria for designation as scientifically novel are as follows:

• *Novel mechanism of action.* This criterion was satisfied if another drug with the same physiological mechanism of action had not been approved for the same indication in either Europe, Japan or the United States more than three years prior to the drug's first approval in either of these three leading markets. Even if another drug with the same mechanism of action had been approved for the same general indication more than three years previously, if that previously approved drug was no longer on the market, this criterion would still be satisfied. In other words, this criterion seeks to exclude drugs for which proof of concept in humans had been demonstrated, to the point of having a drug on the market for the same indication, three years prior to initial approval.

• *Novel structure.* This criterion was satisfied if the drug is the first in a distinctly new family of therapeutic compounds approved in either Europe, Japan or the United States for the general therapeutic indication. As in the case of the new mechanism of action criterion, a three-year grace period applies. In other words, the criterion is still satisfied provided any previously approved drug in the same family has been approved for the same general indication for no more than three years. Even if another drug had been approved more than three years previously, the criterion would still be satisfied if that drug had been withdrawn from market for the general therapeutic indication. The purpose of this criterion is to include drugs that fail the new mechanism of action criterion (for example, because they target a receptor for which drugs have been on the market for more than three years), but nevertheless represent a significant scientific advance because they are the first approved in a distinctly new class of compounds.

Satisfying either criterion would classify the drug as scientifically novel. These criteria were chosen to select drugs whose discovery was not the result of an effort to improve on an already known and at least moderately successful drug. They were also chosen to recognize novel drugs whose development was pursued rapidly well into Phase III trials although efficacy in humans and marketability had not been established.

Classification as scientifically novel is not necessarily an indicator of the degree of scientific insight involved in drug discovery and design. There are many examples of drugs that represent great medicinal chemistry creativity and effort directed at improving upon a drug that had already shown proof-of-concept for a particular mechanism of action. It is also not necessarily an indicator of public health benefit, as there are many drugs that do not meet these criteria but which contribute substantially to health because they are more effective than the first to market compound or they are easier to use (for example, Prezista (2006 pNME) and some of the other second-generation HIV protease inhibitors). Finally, it is not necessarily an indicator of discovery priority of the core compound. For example, other drugs with similar mechanisms of action and basic structures to Viread (2001 pNME), Hepsera (2002 pNME) and Eraxis (2006 sNME) were approved for marketing more than three years before these drugs, yet these three drugs may actually have been discovered first (see below).

The three-year grace period takes into account some of these concerns. It recognizes that modern drug discovery and development often involves a race between several compounds to be the first to obtain marketing approval, and that simply being first-approved-in-class does not necessarily mean that that drug was the first discovered or the sponsoring company took substantially more risks than companies whose drugs were following close behind. Much less does it mean that that drug contributed more to health than subsequently approved drugs in

the class[3]. Note 11 below contains details, presented in a question and answer format, about sources of information and application of these criteria. Applying these two criteria, 118 drugs were classified as scientifically novel and 134 as follow-ons. Table 1 in the main text contains summary statistics for this classification.

Note 5: definitions of universities, biotechnology companies and pharmaceutical companies. The term 'universities' refers not only to universities, but also to academic medical centres such as the Mayo Clinic or Massachusetts General Hospital, government research institutes such as the US National Institutes of Health NIH and the Institute Pasteur in France, and other not-for-profit research centres, such as laboratories funded primarily by the Wellcome Trust or Howard Hughes Foundation.

The term 'biotechnology company' (biotech; B) refers to an independent company formed after 1975 that is focused on science-based drug discovery, and whose involvement in clinical trials and marketing, if any, evolved later from its core focus on science-based discovery.

The term 'pharmaceutical company' (P) refers to a company founded no later than 1975 that devotes considerable effort to discovery, development and regulatory approval of new drugs, as well as marketing of new drugs. If a pharmaceutical company probably had fewer than 1,000 employees when key patents were filed (or key articles published), it was classified as a small pharmaceutical company (indicated as P_S).

The cut-off year for classifying a company as a biotech was chosen in part because this is a natural dividing point between traditional pharmaceutical companies (most of which were formed many years earlier) and the new companies that started to be formed in the late 1970s and 1980s, many of which aimed to exploit new discoveries related to genetic engineering. Also, most of the drugs classified as biotech origin were discovered between the late 1970s and 1990s. Even the oldest and largest biotechnology companies would be considered young at the time they began discovery or development of most of these drugs. However, in the main text section 'Downstream development of biotech drugs', the impact of the largest and oldest biotechnology companies, Genentech, Amgen, Biogen and Genzyme, is considered separately from that of other biotechnology companies.

Note 6: identification of transferees/licensees. University-discovered drugs were classified according to whether the initial development partner was a pharmaceutical company (U>P) or biotech (U>B). The latter were sub-classified according to whether the biotech was in the same region as the university ($U>B_{in}$) or a different region ($U>B_{out}$). For the purpose of this classification, regions are defined as North America, Japan, China, the UK, Continental Europe, Israel, Australia, and other. (If Canada and the US had each been defined as a separate region, the results would not change.) Various sources, principally patents, scientific articles, business articles and US Securities and Exchange Commission (SEC) reports, helped determine which companies first began to develop university-discovered drugs, and approximately how long they carried on their development work. Sometimes the drug was transferred to a second or even a third company before and NDA (or BLA) was filed to the FDA. The classification system focused only on the first transferee. However, that transferee had to have taken significant steps to develop the drug, even though it might have ultimately given up and transferred the drug to another company. Examples of the application of this criterion are given in supplemental information (box) S3, note 2. In the case of university-discovered drugs, the tables in supplementary information S2 identify the university and the first transferee.

As for subsequent transferees, FDA documents identify the NDA applicants and sometimes other licensees around the time of NDA application. Patents, scientific articles, SEC filings and other public information sources were helpful to determine the sequence of licensees from discovery to NDA (or BLA) to marketing. (See also notes 7 and 10 below.)

Note 7: identifying and weighting key contributions to discovery of NMEs and NTBs. For each of the 215 NMEs, investigation was conducted to identify the patents covering the chemical structure of the final compound, patents covering compounds directly preceding discovery of the final compound, and patents covering discoveries that demonstrated therapeutic proof-of-concept for a disease that is a principal target of the final drug, preferably *in vivo*. (In the case of anti-viral drugs where animal models are problematic, positive

responses in standard in vitro assays were accepted as satisfactory indicators of proof-of-concept.) These constituted the key patents for each NME. In the case of drugs that represent new uses of compounds known for a long time, the author sought to identify the persons who showed proof-of-concept of these new uses in living mammals, either through patents claiming these new uses or scientific articles describing these uses.

According to US law (21 USC §355(b) and 31 CFR §314.53(c)), applicants for FDA approval of NMEs must designate the active patents covering the NMEs to ensure that no other company holds rights to the drugs. These patents are listed in the FDA Administrative Correspondence (AC) for most NMEs, which is published online after approval. The author reviewed all these patents, except those clearly covering only formulations or methods of use. Then on the basis of the patent specifications, claims and application dates the author decided which patents likely represented the earliest discoveries of the core therapeutic compound, as well as those that showed it had the desired therapeutic effect in mammals. The active patents for each approved NME listed in the FDA Orange Book were also identified. However, Orange Book patents that were not identified in the FDA AC were only considered if they claimed new compounds or showed first proof-of-concept in living mammals, and were filed prior to the NDA application.

However, if the compound itself has been known for a long time and patents covering it have expired, these usually do not appear in the AC, and definitely not in the Orange Book. The Merck Index[4] typically lists a single patent that covers the active compound. In addition, it also lists key scientific articles describing the compound and its therapeutic use. Thus the Merck Index served as a second source to identify key patents, and as the primary source in the case of 20 drugs where the AC lists no patents covering the active compound. Usually the patents listed in the Merck Index, as well as the key patents listed in the AC, preceded the scientific articles listed in the Merck Index, helping to confirm that the patent inventors deserve credit for discovering the compound. Often the inventors on key patents were also co-authors of the articles describing the compounds listed in the Merck Index.

The patents listed in the Merck Index tend to describe the final formulation of a drug. On the other hand, the AC usually lists more than one patent, with the earliest patent often describing a precursor compound to that described in the Merck Index patents. So, in the 20 cases where the Merck Index was the only source to identify compound patents, these patents were read carefully to see if they cited earlier patents that covered the same basic compound.

Thus the FDA AC, Merck Index, and FDA Orange Book were the main sources for identifying the key patents that covered the core compounds and indicated the drug's principal therapeutic uses. Key patents generally do not include those that, for example, cover only a refinement to the basic compound making it easier to manufacture, store, administer, or absorb, provided the basic compound could be administered safely, could be absorbed sufficiently to obtain a therapeutic effect and could be manufactured in quantities suitable for experimental use. If a modification to the basic compound did, however, significantly increase bioavailability, safety, or even in some cases, ease of manufacture (for example, a breakthrough synthesis method) or administration (for example,. oral as opposed to intravenous), then it was included it among the key patents but generally weighted less than patents covering the basic compound.

Key patents could not be identified for 15 of the 215 NMEs — the sNMEs, Levulan (1999), Abreva (2000), Reminyl (2001), Dacogen (2006), and Veregen (2006); and the pNMEs, Sucraid (1998), Thalomid (1998), Ferrlecit (1999), Panretin (1999), Trisenox (2000), Xyrem (2002), Apokyn (2004), Prialt (2004), Vidaza (2004) and Iplex (2005). All these drugs are based on compounds that have been know for a long time, except Dacogen and Vidaza (which were synthesized in the Czech Academy of Sciences in the 1960s and probably never patented). Prialt (intrathecal therapy for severe pain based on sea snail venom), for which the University of Utah failed to apply for patents and Iplex (which a court ruled infringed patents by Genentech on Increlex, another drug based upon recombinant human insulin-like growth factor-1). For these 15 NMEs, the persons who probably took the key steps to show how the compounds could be used for the intended therapeutic purpose were identified by reviewing the non-compound patents listed in the AC, the scientific articles cited in the Merck Index, and other scientific publications. These articles and their authors served to attribute discovery in the same way as did the key patents and their inventors (see below).

Key patents were grouped into families if they arose from the same application and had nearly identical inventors. Articles that came from the same laboratory (or collaborating laboratories) were similarly grouped into families. When a drug had key patents or articles from unrelated laboratories, the patents and articles were read especially carefully to determine how much weight to attach to each family. In other words, when more than one research group contributed significantly to discovery, each group's proportionate contribution to discovery was estimated, and weights (summing to one) were assigned to each of the key families. As indicated above, the greatest weight was assigned to contributions to discovery of the core therapeutic compound, provided the discoverers had the ultimate therapeutic use in mind, had shown proof-of-concept in a living mammal (or in the case of antivirals, standard assays for antiviral activity), and (in the case of patents claiming a large number of structures) had indicated those particular compounds were serious development candidates. Among compound patents, greatest weight was usually assigned to the earliest patent that claimed the final compound irrespective of enantiomeric conformation (that is, usually the racemic compound). Equal or greater weight was assigned to patents that claim a particular enantiomer only if the racemic compound does not have significant therapeutic activity.

The patent assignees were not assumed to be the inventors' employers, nor were the inventors' addresses on the patents assumed to be the location of their employment. The preferred sources to determine employment at the time of discovery were scientific articles co-authored by the inventors on topics related to the drug and submitted for publication within a few years of the patent application date. (In the absence of other information, the patent application priority filing date was considered to be the date of discovery.) Most biomedical articles give the authors' full names and institutional affiliations (including addresses), and these became the gold standard to determine the type and location of the inventing organizations. The most frequent source for finding such articles was Google Scholar, but the US National Library of Medicine's PubMed databases, the prior art references at the beginning of the US patents, and the Merck Index were also useful starting points. Over 80 percent of the inventors named on all the key patents could be located in a particular laboratory on the basis of articles published within five years of the application priority dates of the patents. Many of these articles dealt specifically with the drug and were co-authored by several co-inventors.

When articles co-authored by particular inventors could not be found, lists of all US patents on which they were named as inventors were obtained. In the case of Japanese-origin drugs, similar lists of Japanese patents were obtained. These lists were inspected for patterns with respect to the invention topics and the identities of the assignees. A large number of patents spanning a long scientific career all assigned to the same company (or to a limited number of companies suggesting movement from one company to another), several of which relate to the drug, are strong evidence of the employer's identity and where the inventor was working at the time of discovery.

In most cases, the address on the patent indicated the location of the laboratory where the inventor was working. However, this is not always the case with some multinational pharmaceutical companies. For example, Pfizer often lists the addresses of inventors working in its Sandwich, UK, laboratory as Groton, Connecticut. Thus, even in cases where an inventor's patent history indicates he/she has been working for the same company, additional efforts were made to obtain scientific articles to indicate the actual place of work.

Occasionally, media reports, or biographical information posted on corporate and university web pages also pinpointed where inventors were working at the time of invention. Using all these sources, the location of employment for most of the remaining inventors was determined with a high degree of certainty.

Fewer than five percent of inventors had such a scant record of publications, patents and media coverage that their place of work at the time of discovery could not be determined. In such cases, if there was nothing unusual about the inventor's address listed on the patent and if the inventor's affiliation was unlikely to significantly alter the drug's attribution of origin, the address on the patent document and the identity of the assignee were assumed to indicate place of work. But in other cases, extra efforts were made to find out information about the inventor. The most important case involved one of two inventors on Enzon's key patent (US

5,951,974) covering PEGylated interferon (Peg-Intron, NTB, 2001). Her address was given as Seoul, Korea. Aside from a few related patents also assigned to Enzon on which she was co-inventor (that also listed Seoul as her address), there were no other patents or articles that shed light on where she was working. A doctoral student in the author's research centre, Sunghee Han, was recruited to research the history of this drug. She kindly found information from Korean sources showing that the inventor was working at Enzon in New Jersey when she made the discoveries covered by the '974 patent, but then returned to Korea.

After determining the place where each inventor worked at the time of discovery, each family of key patents was apportioned according to the type of inventing institution (pharmaceutical company, biotechnology company or university) and also country of invention. In this process, multiple inventors on a single patent were each weighted equally. Then, using the weights described above for each family of key patents, each drug was accorded the value "1" which was apportioned twice, once according to type of inventing institution and once according to national origin. In the case of drugs without key patents, the same process was applied to the families of key scientific articles and their authors. These allocations are shown in the first two tables in supplementary information S2.

During the same ten-year period, the FDA approved 37 NTBs. The FDA does not require applicants seeking approval of NTBs to disclose the covering patents. However, Rader[5] describes the early development history of most of the NTBs and usually indicates the key patents and key transfers of technology underlying each NTB. Thus, the early history of the NTBs was generally more accessible than that of the NMEs. Nevertheless, because many of the NTBs combine distinct technologies, attributing the discovery of NTBs was generally more challenging. Sometimes this required understanding the histories of several core technologies and weighting several families of relevant patents. The absence of a statutorily mandated filter to identify key patents also made attribution more challenging than in the case of NMEs.

Rader's descriptions and the patents he lists served as windows onto the broader discovery history of the NTBs. The patents he cites are important, but they are not as determinative of attribution as are the patents covering NMEs cited in the FDA AC or the Merck Index. Articles by the inventors and articles about the discovery of the drugs (and in one case a direct query to an inventor) played a large role in the attribution of discovery.

The FDA summary documents do not indicate which NTBs approved prior to 2004 were reviewed on a priority basis. Rader, however, states whether the drugs received priority review in most of his case histories. When he does not, notably in the cases of Elitek (2002) and Xolair (2003), it is possible to surmise from his descriptions and various articles about the drugs their review priority. Nevertheless, an element of uncertainty remains. In part for this reason and also because of their small numbers, the NTBs were not analyzed separately according to review priority. At several points in the analysis, however, even the NTBs were classified as priority or standard approved. In such cases, Elitek (2002) and Xolair (2003) were included among the sNMEs and the other NTBs that clearly received standard review. Attributions for the NTBs are shown in the third table in supplementary information S2 along with the organizations that contributed most to the discovery of each drug. Note 12 below presents explanations of how origins were attributed in five complex cases.

For both the NTBs and NMEs, the objective of this study was to obtain an overall picture of the key contributors to the discovery of each drug, and to attribute origins accordingly. Nevertheless, because FDA approval procedures provide a convenient filter for determining inventorship (by requiring NDA applicants to list all active relevant patents), this analysis relies mainly on patents as indicators of inventorship, to the extent that the patent history is compatible with the overall discovery history as revealed by articles and other sources. Indeed, since chance played a role in becoming aware of any discovery history not based upon patents, there is a methodological danger in departing from the attributions suggested by patents alone. However, in the case of a few NMEs, the overall history seemed sufficiently clear to require a different attribution than that suggested by the patents alone. The four cases where such departures most affected the final attributions are presented in Supplementary information S3 (box), note 1.

Throughout this analysis, unless otherwise stated, the attribution numbers that appear represent whole drug equivalent (WDE) values, which represent the summation of the

proportional contributions for a particular category of drugs by a type of organization or country.

Finally, while the goal of attribution accuracy was 10%, the uncertainty range in the case of some of the drugs with mixed attributions is higher. Even in such cases, however, efforts were made to obtain sufficient information so that attributions were accurate to within 20%. In a few complex cases where it was necessary to take into consideration contributions from many discovering organizations, attributions to two decimal places were made, with the realization that attributions to this level are uncertain. Informed comments from readers concerning particular attributions would be welcome by the author.

Note 8: sales analysis. The sales analysis is a parallel analysis to the WDE analysis described above in the preceding note. Sales for individual drugs were allocated among discovering institutions in the same proportions as the drugs themselves. In this way, the attribution of sales according to type of discovering organizations and region could be compared between types of drugs. To make the analysis manageable, peak year sales within the 1999 to 2008 period was selected as the metric. It was assumed that most companies would need two full years to develop their marketing organizations for a new drug and to acquaint the medical community with the drug to the point where the new drug's sales would reflect its comparative market value in relation to other drugs. As 2008 was the last year for which annual sales data was available at the time of analysis, this analysis was limited to the 214 drugs approved by CDER from 1998 through 2005.

Initially, data on 2005 worldwide sales were obtained for the subset of 169 drugs approved between 1998 and 2003. The main data sources for 2005 sales were 10-k, 20-f and annual financial or shareholders reports submitted to the US Securities and Exchange Commission (available under company filings at www.sec.gov). The second most common source was online pharmaceutical business reports. For the three years 2006–2008, annual worldwide sales for most drugs are available at EvaluatePharma (http://www.evaluatepharma.com/Default.aspx). In the case of drugs where generic versions have already been on sale, the EvaluatePharma data includes such generic sales. By these procedures, annual sales for the three or four most recent years were obtained for most drugs. These data were not adjusted for inflation, so as not to give drugs whose sales have declined a boost with respect to drugs whose sales are still increasing.

For the drugs with recent consecutive yearly sales data, if a peak was evident, this was assumed to indicate peak sales and the peak sales year. After the entire analysis was completed, it turned out that the peak sales year was 2008 in about two-thirds of the cases. In other words, two thirds of the 214 drugs approved between 1998 and 2005 still had increasing sales as of 2008.

Where a peak was not evident (that is, where reliable data could not be found or where the trend suggested the peak could be 2005 or earlier) efforts were made to find sales prior to 2005. Again the most common information sources were documents submitted to the SEC or online pharmaceutical business reports. Occasionally medical articles also provided valuable sales data. Care was taken to account for ex-US sales in the case where US and ex-US market rights were divided between different companies. This was not always easy, particularly when ex-US rights were held by a small non-US company that is not obligated to file reports with the SEC. However, it was also difficult in the case of some low-sales drugs marketed by large pharmaceutical companies, because the large companies often do not report sales of their low-sales drugs to the SEC.

Considerable uncertainty regarding the peak sales estimates remains for three drugs whose likely peak sales exceed $50 million. These are: Septocaine (2000 sNME) a dental anaesthetic long used in Europe that is used increasingly frequently in the United States, TNKase (2000 NTB) a thrombolytic for acute myocardial infarction, and Elitek (2002 NTB), a recombinant urate oxidase for hyperuricemia due to tumour lysis in chemotherapy patients. In the case of TNKase, the peak sales estimate of $260 million in 2008 was pieced together from information from Genentech's and Roche's SEC filings. The estimated $350 million peak year sales in 2008 for Septocaine assumes a $1 manufacturer's wholesale bulk price per cartridge and 200–300 million cartridge sales annually in the US. The peak year sales estimate for Elitek of $100 million in 2008 is based upon data for 2001–2005 kindly provided by EvaluatePharma showing a significant rise in annual sales from 2003 to 2005 (with sales reaching $57 million

in 2005), as well as medical articles indicating increasing use of Elitek by North American oncologists, to the point that some consider it the preferred therapy for tumour lysis syndrome (Ikeda 2008).

Peak year estimates are also quite uncertain for eleven other drugs. However, estimates were calculated based on a variety of data sources, including clinical medical articles providing data on the number of patients who might take the drug, and articles describing competing medicines. About ten of these probably have peak sales no greater than $20 million. The eleventh, Mifeprex (2000 pNME), was estimated to have recent annual sales of about $40 million assuming that about 200,000 persons used the drug in the US in 2008 and (very roughly) that the wholesale price charged by whatever companies are approved to market the drug is about $50 per dose and the US accounts for about a quarter of world sales. Peak sales with considerable uncertainty (i.e., uncertainty ranges from about –75% to +100%, or + $10 million, whichever is greater) are indicated in red in the tables in supplementary information S2, while peak sales with uncertain ranges from about +30% are indicate in blue in the same tables. Likewise, a peak year indicated in red means that the actual peak year could be five or more years away, while a peak year indicated in blue means that the actual peak year probably is within 3 years of the indicated year.

For each drug approved between 1998 and 2005, peak sales were attributed among the inventing organizations and countries by multiplying the sales figure by the attribution proportions shown in the tables in supplementary information S2. As in the case of the WDE analysis, sales attributions for individual drugs were summed to calculate total attribution of sales according to inventing organization or country. Because this analysis is so heavily influenced by blockbuster drugs with peak sales over $1 billion, errors of even 100% in the estimate of peak sales less than $50 million will not significantly affect the analysis.

Note 9: orphan drugs. This study also identified the companies that have been instrumental in the post-discovery commercialization of each drug — the first companies to have pursued discovery or development concertedly, the NDA applicants, and the companies that first marketed each drug in the US.

Drugs that are designated as orphans in the US, Japan or Europe were also analyzed separately to investigate whether orphan drug legislation encourages indigenous discovery of drugs responding to niche medical needs.

Lists of orphan indications approved by the FDA for various drugs are available at http://www.fda.gov/orphan/designat/list.htm. Similar lists are available for orphan indications approved by Japanese and European regulatory authorities (Supplementary information S4 (box), note 4). In order to exclude from the analysis drugs that have been approved for non-orphan indications, both lists were cross checked against FDA approved indications. Also the Japanese list was cross checked against lists of Japanese approvals for non-orhpan indications.

Note 10: Figure 3. Figure 3 is based on the data in the '*main attribution code, applicant code* and *marketer code* columns in the tables in supplementary information S2. The main attribution code identifies the type of organization to which a majority of the discovery of a particular drug is attributed. These codes are listed at the bottom of the supplementary information S2 tables. When two types of organizations contributed roughly half to discovery, two majority attribution codes were assigned and attribution was half to each type. A similar division of development milestones was carried out in rare cases when two companies applied for FDA approval (for example, Lilly and Icos jointly filing the NDA for Cialis) and in the more common cases of two or more companies undertaking initial marketing in the US. Fig. 3 aggregates the majority attributions at each stage for each type of organization. For the sake of simplicity, Ps and P* drugs are grouped together and designated by light blue, but U→P_s and U→P* are grouped with other U→P drugs and colour-coded green. Unlike in Figures 1 and 2, U→B_{in} drugs are not grouped together with biotech discovered drugs, but instead all U→B drugs are grouped together and colour-coded pink in the discovery column.

The Figure 3 columns for 'FDA applicant' and 'First year marketing in US' simply sum the entries for biotechnology companies and pharmaceutical companies as a whole from the applicant and code columns in the supplementary information S2 tables. They indicate the aggregate activity of biotechnology companies and pharmaceutical companies at these stages,

without regard to whether they were the original discoverers or licensees and, if the latter, whether the licensor was a biotechnology company or a pharmaceutical company. However, the analysis in the text is based in part on these considerations.

Note 11: Explanations in Q&A format on criteria for scientifically novelty.

Q1. What about cases when the discovery of a drug that eventually became the first in its class to be marketed, was coincident with (or even later than) the discovery of the core compound that eventually became a later drug in the same class?

A1. This may be the case with some nucleoside/nucleotide reverse transcriptase inhibitors for HIV and HBV. Lamivudine (Epivir) was the first drug in this class approved by the FDA for HIV in 1995 and for HBV in 1998. It was probably invented by researchers at IAF Biochem International in Canada about 1989. The core compounds of tenofovir (Viread, approved as a pNME in 2001 for HIV) and adefovir (Hepsera, pNME approved 2002 for HBV) were discovered by researchers at the Czech Academy of Sciences around 1986. Researchers at Gilead and BMS subsequently developed prodrugs in the 1990s. Similarly, the lead compound for the third-approved echnocandin antifungal, anidulafungin (Eraxis 2006 sNME), was discovered in 1974, probably before the compounds that became the first and second approved echinocandin antifungals, caspofungin (Cancidas 2001 pNME) and micafungin (Mycamine pNME approved in Japan in 2002 and in the US in 2005). Nevertheless, this study does not classify tenofovir, adefovir, anidulafungin and drugs with similar histories to be scientifically novel drugs, in part because it is possible that the compounds that became lamivudine, caspfungin and micafungin were themselves discovered even earlier. Another reason is the criteria for scientific novelty are intended not only to select compounds that were discovered early, but also to recognize the efforts of companies that follow through with development. Even though the basic compounds that became tenofovir, adefovir and anidulafungin may have been discovered earlier than those that became the first-in-class approved drugs, it is possible, perhaps likely, that the companies that ultimately developed them waited until proof-of-concept had been shown by other drugs to push forward with development.

In contrast, ketotifen (Zaditor, pNME approved in 1999 for allergic conjunctivitis) was classified as having a new mechanism of action (mast cell stabilization, H1 receptor blockade and anti-eosinophil chemotaxis) even though earlier approved drugs, for example, emedastine (Emadine sNME 1997) and olopatadine (Patenol, sNME 1996) have some of the same MAs, and Novartis's decision to seek FDA approval was probably motivated by olopatadine's success in the US market. This classification is based in part on the fact that the key patents covering ketotifen date from the early 1970s (US 3682930 issued to Sandoz in 1972) while those covering olopatadine and emedastine date from the 1980s (respectively, EP 235796 issued in 1987 to Kyowa Kakko, and US 4430343 issued in 1984 to Kanebo), but more importantly on the fact that Sandoz began marketing ketotifen in Europe in the 1970s before olopatadine or emedastine were marketed in Japan or elsewhere.

Q2. Why was three years chosen as a grace period, and how did this influence actual classifications?

A2. The author made the assumption that approval within about three years of the first-in-class approved drug indicates parallel discovery and development — that the commitment to push for approval had been made before efficacy, safety and market demand with respect to first-in-class drug were clear. Thus erlotinib (Tarceva) was classified as a drug with a new mechanism of action, even though gefitinib (Iressa) was approved one year earlier in 2003. Similarly, 5-aza-deoxycytidine (Dacogen, sNME approved 2006 for myelodysplastic syndrome (MDS)) was considered to have a new mechanism of action, because it was approved within two years of its congener, 5-azacytidine (Vidaza, pNME approved 2004 also for MDS).

Q3. What were the most frequent information sources to make this determination?

A3. 'Fresh From the Pipeline' articles in *Nature Reviews Drug Discovery*, and articles from the *Journal of Medicinal Chemistry* and other pharmaceutical journals found by searching on the name of the drug + mechanism of action or discovery.

Q 4. What are the principal qualifications with respect to applying these criteria?

A4. Pro-drugs of previously marketed drugs were generally not considered to have new mechanisms of action or structures.

Also, if the structure and biochemical mechanism of action are similar to those of a drug from the same company previously approved for a different therapeutic use, the new drug was not considered to be novel if the first approved drug can be used for the indication for which the new drug is approved. Thus, Lucentis (NTB 2007 from Genentech) was not considered to be novel in view of Avastin's approval in 2004 and use for macular degeneration. Similarly, the following were not consider as novel: alosetron (Lotronex pNME 2000 from Glaxo) in view of ondansetron's 1991 approval and use for diarrhea-associated irritable bowel syndrome, lenalidomide (pNME 2005 from Celgene) in view of thalomide's1998 approval and use for multiple myeloma, and temsirolimus (pNME 2007 from Wyeth) in view of sirolimus's 1999 approval and use against various cancers.

Q5. What are some examples of the interaction between the two criteria?

A5. Although ixabepilone (Ixempra, 2007 pNME) probably binds to the same pharmacophore on microtubulins as do taxanes, and its basic anticancer mechanism is the same, that is, stabilization of microtubules which interferes with mitosis, ixabepilone is the first approved drug in a new class of compounds, the epothilones. The different structure of these compounds probably accounts for binding properties different from those of the taxanes, and probably also for their different resistance profiles[6]. Similarly, the non peptide HIV protease inhibitor, Aptivus (2005 pNME) was considered novel in relation to previously approved peptide based protease inhibitors in view of its novel non-peptide-based structure.

On the other hand, the anti-cancer drug, Alimta (2004 pNME), was not considered to have a new mechanism of action in view of its similarities to methotrexate and raltitrexed (not approved in the US but sold in Europe as Tomudex) even though its activity profile against enzymes involved in the synthesis of purines and pyrimidines differs from that of the previously approved drugs. This is because of the structural similarities between Alimta and methotrexate and raltitrexed, the fact that the anti-neoplastic effect of all three drugs is based upon inhibiting folate dependent pathways, and the fact that each of the enzymes Alimta inhibits, thymidine synthase, glycinamide ribonucleotide formyltransferase, and dihydrofolate reductase, are also inhibited by either methotrexate or raltitrexed[7].

As for three pNMEs approved for glaucoma (unoprostone (Rescula 2000) bimatoprost (Lumigan, 2001) and travoprost (Travatan 2001)), although unoprostone is sometimes classed as a docosanoid or prostanoid, and bimatoprost sometimes as a prostamide (both in distinction to prostaglandins) all have prostaglandin-like structures. Also, the physiological mechanism of action of all three of these drugs appears to be similar; that is, all decrease intra-ocular pressure by increasing the outflow of aqueous fluid — probably by relaxing the trabecular meshwork in the outflow pathway. In this respect, the mechanism of action and structure are similar to that of the prostaglandin analogue, latanoprost (Xalaten pNME), approved in 1996 for glaucoma[8-10]. Thus none of these three glaucoma drugs were considered to have a new mechanism of action or to be first in a new structural class.

Q6. What were some of the other challenging cases in applying these criteria?

A6. Drugs ultimately classified as follow-ons: Refludan (1998 pNME), Solage (1999 sNME), Angiomax (2000 sNME), Zonegran (2000 sNME), Abilify (2002 sNME), Bystolic (2007 sNME).

Drugs ultimately classified as scientifically novel: Aggrastat (1998 pNME), Integrilin (1999 pNME), Abreva (2000 sNME), Elidel (2001 sNME), Reminyl (2001 sNME), Elitek (2002 NTB) and Amitiza (2006, sNME).

Insights are welcome from readers who are familiar with the discovery histories of these or other drugs. The author will respond to informed queries as to the rationale for classifying particular drugs as follow-on or novel.

Note 12: Examples of attribution of origins in five complex cases.

Case 1. Gemtuzumab ozogamicin (Mylotarg, approved 2000) is a guided-warhead anti-cancer drug consisting of an anticancer drug (ozogamicin) attached to a monoclonal antibody (gemtuzumab) designed to bind to myelocytic leukemia cells. Researchers in Lederle's (now Wyeth's) Pearl River, New York, laboratory designed the anticancer drug and the means to attach it to the antibody and were responsible for the overall design and testing of the drug. Lederle/Wyeth researchers received the 2004 Discoverers Award from the Pharmaceutical Research and Manufacturers of America (PhRMA) for their development of Mylotarg—the first such warhead-tipped antibody cancer drug approved by the FDA. (It is not clear why Mylotarg was approved as a pNME rather than an NTB.) Their discoveries related to the anticancer drug are represented by three independent patents listed in the FDA Administrative Correspondence (AC). Their discoveries related to linking the anti-cancer drug to the antibody are covered by three other patents listed in the AC, all of which originate from the same application. According to Rader (2006), the antibody to the CD33 receptor (which tends to be expressed uniquely on the surface of myelocytic leukemia cells) originally came from the Frederick Hutchinson Cancer Research Center in Seattle, but CellTech in Berkshire, England, took the lead in humanizing it using its own technology and technology in-licensed from Protein Design Laboratories (PDL) of Mountain View, California. In addition to the six patents assigned to Wyeth, the AC lists two patents assigned to PDL. However, these are not specific for an anti-CD33 antibody. The AC lists no patents assigned to CellTech. However, an article[11] co-authored by ten Wyeth researchers, one CellTech researcher and two Fred Hutchinson researchers describes the development of Mylotarg.

One of the four inventors on Lederle's first patent (US 4970198) covering the anticancer drug, David Lebeda, was working at the US Department of Agriculture's Northern Regional Research Center in Peoria, Illinois, at the time of the patent application in the mid-1980s. Lederle scientists obtained the microorganism that was found to produce calicheamicins, the family of anti-cancer compounds from which they derived ozogamicin, from the culture collection where Lebeda worked. Thus Lebeda was assumed to have played a key role in providing the microorganism. However, he is not an author on any publications in the PubMed data base that are related to the Mylotarg or its components. Thus, he was assumed not to have played a major role in elucidating the anticancer properties of calicheamicins or in deriving ozogamicin from them. As for the other inventors listed on the Lederle patents, all were confirmed as being employed by Lederle/Wyeth on the basis of publications or a clear history of assigning patents to Lederle, Wyeth, or American Cyanamid (Lederle's parent until the name was changed to Wyeth).

In attributing discovery to the various laboratories, inventing/developing the warhead, the linker, and the antibody, and integrating all of these into a prototype drug, were weighted 0.25 each. Taking all the above factors into account, the warhead, linker and integration were attributed to Lederle. Thus 75% of the discovery was attributed to US pharmaceutical companies. Of the remaining 25% attributable to the antibody, 15% was attributed to CellTech researchers (UK biotechnology company) because they seem to have played the main role in humanizing the antibody. 5% was attributed to PDL (US biotechnology company) because CellTech relied on PDL's technology for humanizing antibodies and, even though PDL's technology probably was not specific to CD33 receptors, Wyeth evidently thought that PDL's patents covered the antibody when it submitted its application to the FDA. Finally, 5% was attributed to Fred Hutchinson (US university transferring to out-of-region biotechnology company) whose researchers developed the original mouse antibody against CD33. Lebeda's/USDA's contribution was ignored for the reasons indicated above.

Cases 2 and 3. Remicade and Humira are both monoclonal antibodies against tumour necrosis factor (TNF), which is implicated in the inflammation that causes rheumatoid arthritis (RA) and other autoimmune diseases. Remicade is a recombinant, chimeric (partially humanized) mouse antibody approved in 1998 to treat Crohn's disease and in 1999 to treat RA, while Humira is a recombinant, fully human antibody approved in 2002 for RA.

In the case of Remicade, at least ten US patents with the same two New York University (NYU) and same four Centocor inventors claim chimeric anti-TNF antibodies or describe the use of such antibodies to treat RA. These patents all trace their origins to the same application filed in March, 1991. A list of all 32 US patents assigned to NYU that have Jan Vilcek as one of the inventors (Jan Vilcek and Junming Le are the two NYU inventors whose names appear consistently on these patents), indicates that probably all of these 32 inventions that deal with TNF antibodies as RA therapy had Centocor co-inventors (most often, the same four Centocor inventors). A 1995 article showing that chimeric TNF antibodies protect mice from TNF-mediated diseases was co-authored by seven Centocor and two NYU researchers (including the same six who are co-inventors on so many related patents) as well as one researcher at the Hellenic Pasteur Institute in Athens[12]. In other words, the patent and publication record suggests a long history of cooperation between Centocor and NYU related to Remicade, along with consistency in terms of the individual researchers involved in this collaboration. Rader's[5] history indicates Remicade was developed by Centocor in collaboration with NYU researchers. Centocor, based in Malvern, Pennsylvania, was an independent biotech until 1999 when it was bought by Johnson & Johnson in 1999. Bearing in mind, these factors, the discovery and early stage development of Remicade was attributed 65% to Centocor (biotechnology company) and 35% to NYU (university transferring to same region biotechnology company), and 100% to the US in terms of national origin.

Humira arose from collaboration between Knoll Pharmaceuticals and Cambridge Antibody Technology (CAT) in the UK[5]. CAT developed and provided the phage display technology that enabled creation of a fully human anti-TNF monoclonal antibody. Rader[5] lists US 6090382 and 6258562 as the two key patents covering Humira. These patents originate from the same application filed in February 1996. Both have the same twelve inventors. Six were working at BASF/Knoll's research centre in Worcester, Massachusetts, from the mid 1990s. Six were employees of CAT around the same time, although one of these, Hendricus Hoogenboom, had in 1995 become an associate professor in the University of Maastricht in the Netherlands, and some of the CAT inventors had ties with Cambridge University (CAT being a Cambridge start-up company). In the 1990s, Knoll Pharmaceuticals was BASF's US pharmaceutical research laboratory. BASF sold its pharmaceutical operations, including its Massachusetts research centre, to Abbott in 2001. Abbott sponsored Humira's application for approval by the FDA and markets Humira worldwide, except for Japan where it co-markets with Eisai.

However, PepTech, a UK-Australian biotech company founded in 1985, has patents covering monoclonal antibodies to TNF (for example, US 5644034 and 6593458, many of which originate from Australian patent applications filed in 1989). Knoll licensed these antibodies for use in Humira. The extent to which Knoll actually relied on PepTech's discoveries is not clear. PepTech's research on TNF antibodies dates from 1988 (REF. 5). Articles on this topic by Peptech scientists began to be published around 1991 (see REF. 13) and the first of Peptech's US TNF antibody patents (no. 5644034) was issued in November 1994, nearly 1.5 years before BASF applied for its Humira patents. Therefore, Peptech's TNF antibody technology was assumed to have been used by Knoll to develop its TNF antibodies. One of the two inventors of Peptech's US patents, Roger Aston, served as CEO of CAT before becoming Peptech's CEO in 1995 and had other close links with the UK, reinforcing the likelihood that Peptech's technology played a role in the discovery of Humira. However, the balance of evidence suggests that Roger Aston's work on TNF antibodies was centred at Peptech in Australia. Taking these factors into consideration, 40% of Humira's discovery was attributed to CAT (and all its inventors were assumed to have done their relevant work in CAT's UK laboratory), 40% to Knoll (and all its inventors were assumed to have done their relevant work in the US), and 20% to Peptech (and all its inventors were assumed to have done their relevant work in Australia). Thus the final attribution, in terms of type of

discovering organization was 0.6 biotechnology company and 0.4 pharmaceutical company, and in terms of national origin, 0.4 UK, 0.4 US, and 0.2 Australia.

Centocor also licensed Peptech's TNF antibody patents. In 2002, Centocor's new parent, Johnson & Johnson, decided Remicade was not infringing these patents and stopped royalty payments, as did Abbott in the case of Humira. The dispute with Abbott was resolved quickly, largely in Peptech's favour[5]. However, the dispute with Johnson and Johnson was not resolved until the end of 2004, and although it involved some payments to Peptech, the details are confidential and the overall balance of the settlement is not clear. The academic articles co-written by NYU and Centocor describe their own methods for making TNF antibodies and appear to make no reference to articles by Peptech researchers (see REF. 14). Although this may have been deliberate, it nevertheless suggests that, to the extent the Centocor and NYU researchers made use of Peptech's discoveries, they made use of knowledge that was publicly available, regardless of whether the techniques were protected by Peptech's patents. Also Centocor/NYU's work was earlier in time compared to Knoll/CAT's. Furthermore, in Roger Aston, there was a clear human link between Peptech and CAT, that likely directly informed CAT's cooperation with Knoll to develop the anti-TNF antibody that became Humira. For these reasons, although some of Humira's origins were attributed to Peptech, none of Remicade's were.

For similar reasons, none of Remicade's or Humira's origins were attributed to the Weizmann Institute in Israel, which also invented mouse TNF antibodies. (See US Patent 6090923 issued in 2000 claiming priority to applications filed in Israel in 1984.) Through a research alliance with Weizmann, the Swiss pharmaceutical company Serono obtained control over these patents. In 2000, Serono issued licenses to Centocor and Knoll covering Remicade and Humira, respectively (Centocor's license was part of a litigation settlement between Centocor and Serono). Terms are not public. However, Centocor does not pay royalties on sales, while Knoll does[5]. A comparison of the US patents issued to Yeda Research and Development (Weizmann's technology commercialization arm) and Peptech suggests that the latter's patent coverage of TNF antibody technology is more extensive than the former's. Thus it was assumed that while Peptech's technology directly contributed to Humira, Weizmann's either did not or it was small compared to Peptech's.

Cases 4 and 5: Pegylated interferon alpha. Peg-Intron, approved in 2001 and marketed mainly by Schering-Plough, and Pegasys, approved in 2002 and marketed mainly by Roche, are different subspecies of recombinant interferon alpha that are linked to polyethylene glycol (PEG) molecules and are useful for treating some viral infections and cancers. These linked PEG molecules prolong the time the interferon remains active in the body and thus increase its effectiveness and lowers the total required dose. Methods for making recombinant interferon alpha, and the recombinant products themselves, were claimed in competing patents filed by Genentech and Biogen in 1980 and 1981. Eventually, Roche emerged with rights to the 2a subspecies derived mainly from Genentech's patents and Schering Plough emerged with rights to the 2b subspecies derived mainly from Biogen's patents, and these were approved as Intron A and Roferon A, respectively, both in 1986 (REF. 5). Thus, the novelty of the interferon alpha NTBs in the study sample derives from their PEG linkages. The largest demand for both these PEGylated interferon alphas is weekly injections to treat chronic hepatitis C.

The PEG technology used in Peg-Intron was developed by Enzon in Piscataway, New Jersey. Enzon, was formed in 1982 by researchers at Rutgers University (see REF. 15 and US patent 4179337 filed in 1977). As of April 2007, Enzon had received 111 US patents and expanded its expertise in this field. Its focus continues to be PEG-linked drugs that have improved characteristics over the originals. US patent 5951974 appears to be the first of Enzon's patents claiming a pegylated form of interferon, and it is the key patent designated in Rader's[5] review of this drug's history. All of the inventors were researchers at Enzon at the time the initial application was made in 1993. The fact that Enzon continued to manufacture PEG-Intron in the years immediately following its approval implies that most of the input into the early development of PEG-Intron should be attributed to Enzon[15]. However, Rader[5] notes that the single chain PEG molecule attached to interferon alpha in PEG-Intron is first-generation PEG technology that Enzon no longer employs in the newer products it is developing. Thus Enzon's reliance on the original Rutgers invention to create Peg-Intron

probably was not de minimis. As for Schering-Plough, a review of its interferon patents indicates that the only patents pertaining to pegylated interferon cover methods of treatment (e.g. US 5908621) or improved formulations for storage (US 6180096). Therefore, the final attribution for PEG-Intron is 0.85 to Enzon and 0.15 to Rutgers, or 0.85 biotechnology company, 0.15 university (transfer to same region biotechnology company), with national attribution being entirely to the US.

In the case of Pegasys, approved one year after PEG-Intron, PEGylation technology came from Shearwater Pharmaceuticals (now Nektar Therapeutics). Shearwater was formed in 1992 by Dr J. Milton Harris of the University of Alabama to develop a variation of pegylation technology that may have overlapped with Enzon's. The initial patent covering this technology (US 5252714) named Dr Harris and another University of Alabama researcher as inventors. It was filed in 1990 and assigned to the University of Alabama. Roche initially tried to develop its own PEGylated version of Roferon A, but turned to Shearwater when its molecule failed to show substantial improvement over non-PEGylated Roferon A. However, unlike Enzon's patents, Shearwater's patents appear to cover only the PEG molecules, not pegylated therapeutic drugs. There is a suggestion in Rader (2006) that Roche played the main role in joining Roferon A and Shearwater's PEG molecule. US patent 7201897 issued in April 2007, based upon patent applications made in 1996 and 1997, seems to confirm this — the two inventors were working in Roche's Nutley, New Jersey research centre in 1996. The summary in the patent body states, "the invention is a new class of PEG derivatives of interferon alpha. The conjugate of this invention has a branched PEG structure [a characteristic of the Shearwater molecule], …thus doubling the attached PEG mass without multiple sites of pegylation." This branched pegylated interferon appears to have slightly better clinical efficacy than the single-branch-PEGylated PEG-Intron. Taking into account the importance of Shearwater's PEG molecule, but also Roche's role in synthesizing and improving the final conjugate, the discovery of Pegasys was attributed half to Roche, and half to Shearwater and the University of Alabama. As for the latter's contribution, considering that Dr. Harris remained on the faculty until 2000 and only six years elapsed from the time of the filing of Dr. Harris's first PEG patent to the time Roche applied for a patent on a conjugate that resembled the final drug, the University's direct contribution is not de minimis. On the other hand, Shearwater's 26 issued US patents indicate the company was continually improving the original discovery. Pegasys incorporates a branched, second-generation molecule indicating that Shearwater is probably the source of this molecule. In view of these factors, 10% of Pegasys's discovery was attributed to the University of Alabama, and 40% to Shearwater. Thus the final attribution was 0.5 pharma, 0.4 biotech and 0.1 university (transferring to same-region biotechnology company). National origin was attributed 100% to the US — all of Roche's inventors on its Pegasys related patents were working in its Nutley, New Jersey, laboratory.

Note: The fact that a particular drug relies on in-licensed technology does not necessarily imply that the supplier of that technology ought be considered a co-discoverer. In particularly, the inventors of broadly useful enabling technologies incorporated in several drugs are not considered to be co-discoverers of these drugs. For example, the NTB Enbrel incorporates immunoadhesin (anti-body like fusion protein) technology in-licensed from Genentech and also co-transformation (genetic engineering) technology in-licensed from Columbia University. However, because these technologies are used in many drugs, none of Enbrel's origins are attributed to Genentech or Columbia.

References

1. DiMasi, J. A. & Grabowski, H. G. The cost of biopharmaceutical R&D: Is biotech different? *Manag.Decision Econ.***28**, 469–479 (2007).
2. Crabtree, P.. New drug price tag: $1.2 billion—Biotechs criticized for rising treatment costs. *San Diego Union-Tribune.* 10 Nov (2006).
3. Schwieterman W. D. Regulating biopharmaceuticals under CDER versus CBER: an insider's perspective. *Drug Discov Today* **11**, 945–951 (2006).

3. DiMasi, J. A. & Paquette, C. The economics of follow-on drug research and development. *Pharmacoeconomics* **22** Suppl 2, 1–14 (2004).
4. O'Neil, Maryadele J. (Ed.). 2006. *The Merck Index: An Encyclopedia of Chemicals, Drugs, and Biologicals.* (Merck & Co., Inc., Whitehouse Station, NJ, USA, 2006).
5. Rader, R. A. *Biopharmaceutical Products in the U.S. and European Markets* 6th edition, Volumes 1 & 2 (Bioplan Associates, Rockville, Maryland, 2007).
6. Lee, F. *et al.* Preclinical discovery of ixabepilone, a highly active antineoplastic agent. *Cancer Chemother. Pharmacol.* **63**, 157–166 (2008)
7. Adjei, A. Pemetrexed (ALIMTA), a novel multitargeted antineoplastic agent. *Clin. Cancer Res.* **10**, 4276s (2004)
8. Lim, K. *et al.* Mechanism of action of bimatoprost, latanoprost, and travoprost in healthy subjects, a crossover study. *Ophthalmology* **115**, 790-795 (2008).
9. Susanna, R *et al.* Current status of prostaglandin therapy: latanoprost and unoprostone. *Surv. Ophthalmol.* **47**, (suppl 1.) S97–S104 (2002).
10. Woodward, D. F. *et al.* The pharmacology of bimatoprost (Lumigan™). *Surv. Ophthalmol.* **45,** (suppl 4) S337–S345 (2001).
11. Hamann, P. R. *et al.* Gemtuzumab ozogamicin, a potent and selective anti-CD33 antibody-calicheamicin conjugate for treatment of acute myeloid leukemia. *Bioconjugate Chem.* **13**, 47–58 (2002).
12. Siegel, S. A. *et al.* The Mouse/human chimeric monoclonal antibody cA2 neutralizes TNF in vitro and protects transgenic mice from cachexia and TNF lethality *in vivo*. *Cytokine* **7**, 15–25 (1995).
13. Rathjen, D. A. *et al.* Antigenic structure of human tumour necrosis factor: recognition of distinct regions of TNF alpha by different tumour cell receptors. *Mol. Immunol.* **28**, 79–86 (1991).
14. Knight, D. M. *et al.* Construction and initial characterization of a mouse-human chimeric anti-TNF antibody. *Mol. Immunol.* **30**, 1443–1453 (1993).
15. Katsnelson, A. Abe Abuchowski. *Nature Biotechnol.* **25**, 1050 (2006).