self-sustaining and does not rely on outside funding.

The most popular plasmids in the collection are empty backbones created for specific gene expression or knockdown experiments, control plasmids, and constructs used for generating lentiviruses and retroviruses. A quick look at Addgene's most requested plasmids, according to laboratory (**Table 1**), reveals a collection of vectors that can be used in various applications across multiple disciplines. If a BRC like Addgene were not archiving and distributing these valuable reagents, they would be far less accessible to the scientific community[3]. Indeed, many researchers, especially those outside the discipline of the contributing laboratory, might not even realize that some of these powerful tools exist. Addgene has become a global repository, sending out approximately half of its requests to scientists outside the United States.

Addgene now distributes genomic resources for large-scale projects, such as the Zinc Finger Consortium (http://www.zincfingers.org/), the Structural Genomics Consortium (http://www.thesgc.org/) and the Center for Genomic Engineering (http://www.cge.umn.edu/). Moving forward, Addgene hopes to collaborate with additional groups to help support their archival and distribution efforts.

In addition to archiving and distributing a physical reagent, Addgene also plays a crucial role by archiving information about these reagents and making it accessible to all potential users through an online database. Addgene's website receives an average of 35,000 page views per weekday. Having clone information available helps with reproducibility and future use, especially because checking the accuracy of this information is often an onerous task for many laboratories. Similar to other BRCs, Addgene can handle large volumes of samples and data, which facilitates the development of efficient, large-scale processes for standardizing quality control and maintaining comprehensive databases of information. Currently, Addgene sequences key regions of all incoming constructs, which helps maintain a standardized bar for accuracy throughout the repository.

Addgene has developed one of the first electronic material transfer agreement (MTA) systems, which has helped expedite the MTA process. Over the past few decades, there has been an increase in the use of MTAs for transferring reagents between academic and nonprofit organizations. Although MTAs may be a practical means of maintaining control of reagents by the institution, they can often cause long delays for the researcher looking to obtain these reagents. Addgene has streamlined the technology transfer process by (i) using the universal biological material transfer agreement (UBMTA) as the basis for all transfers, (ii) making the agreements as consistent as possible across all institutions and (iii) allowing for electronic signatures from institutions that both contribute and request materials. This system has been used for >80,000 orders from >2,500 institutions worldwide. As more technology transfer offices have adapted to this system, the time required for MTA approval has been halved, with the median time now <36 h. Moving forward, it would be more efficient for institutions to implement a similar electronic MTA system for all academic resource transfers.

Ultimately, BRCs like Addgene will be important for guiding academic laboratories into a new age of high-throughput research and corporate funding. We are seeing a paradigm shift in the pharmaceutical industry toward greater collaborations with academia in early stages of the drug development pipeline. With this shift, institutions will become more reliant on BRCs to help manage large collections and streamline legal paperwork, while still promoting open access to the research community.

ACKNOWLEDGMENTS

We would like to thank our community of >1,100 Addgene depositors (http://www.addgene.org/browse/pi/) for their helpful discussions on resource-sharing practices over the years.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturebiotechnology/.

Melanie Herscovitch, Eric Perkins, Andy Baltus & Melina Fan

Addgene, Cambridge, Massachusetts, USA.
e-mail: melanie@addgene.org

1. Weaver, T., Maurer, J. & Hayashizaki, Y. *Nat. Rev. Genet.* **5**, 861–866 (2004).
2. Fan, M., Tsai, J., Chen, B., Fan, K. & LaBaer, J. *Science* **307**, 1877 (2005).
3. Campbell, E.G. *et al. J. Am. Med. Assoc.* **287**, 473–480 (2002).

Use of genome-wide association studies for drug repositioning

To the Editor:

Over the past few years, large investments have been made in genome-wide association studies (GWAS) with the expectation that some of these studies would lead to the identification of novel therapeutic modalities or allow selection of patients who would respond better to therapeutic interventions. Although the results have provided valuable biological insights for many common diseases, the translation of the genetics findings from GWAS into the clinic remains limited and a topic of intense debate. Among the factors that could explain this situation are that the road from a gene target to an approved marketed drug takes in general more than ten years and most GWAS results have only been obtained over the past four years. Furthermore, because the effect size of the common variants identified by GWAS, alone or in aggregation, is generally modest, the impact in terms of personalized, individually tailored medicine has been negligible. We present here an analysis of another potential application of GWAS data—drug repositioning. In the following study, we assess the utility of GWAS in systematically and rapidly identifying alternative or refined indications for existing drugs.

The complete analysis of our workflow is described in **Figure 1**. Our approach began with the construction of a list of GWAS genes associated with disease traits. We used the catalog of published GWAS data from the US National Human Genome Research Institute (NHGRI; Bethesda, MD; http://www.genome.gov/gwasstudies). This resource contains an exhaustive description of trait/disease-associated single nucleotide polymorphisms (SNPs). At the time of our analysis (February 14, 2011) the GWAS catalog contained 796 publications with 4,818 rows of data, each row corresponding to an association between a trait and an index SNP. In an attempt to minimize the inclusion of false-positive signals, we eliminated associations annotated as not replicated, and with $P > 1e^{-7}$ (**Supplementary Methods**). An additional 400 associations listed in **Supplementary Table 1** were excluded because the associated traits were anthropometric and not relevant in the drug discovery context of our analysis. The remaining 1,515 rows from 361 publications

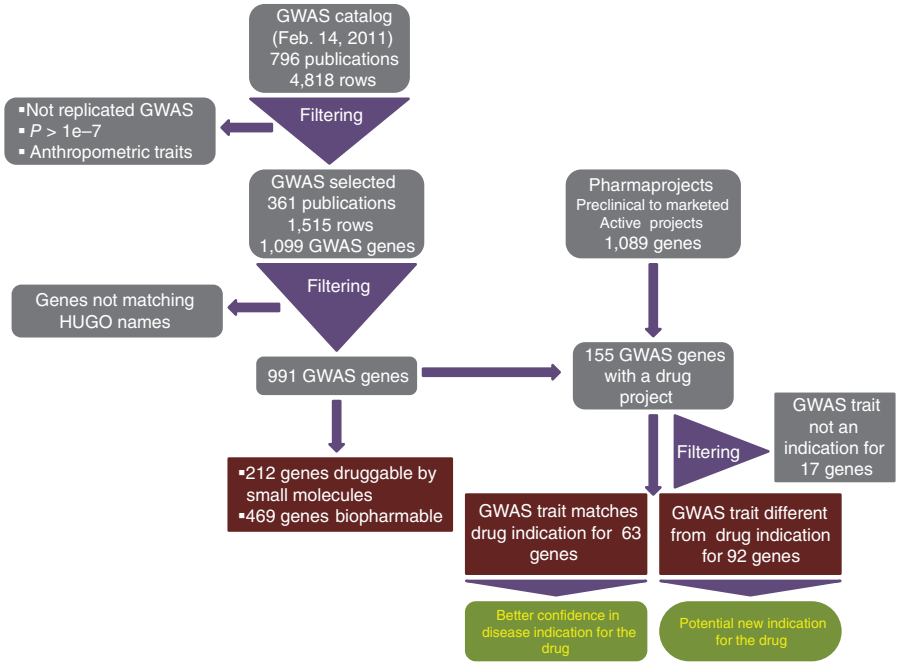

Figure 1 Analysis pipeline. Nine hundred ninety-one GWAS-associated genes were selected from the GWAS catalog after two filtering steps (**Supplementary Methods**). These genes were evaluated as potential drug targets for small molecules and biopharmaceuticals. One hundred fifty-five of these 991 genes were also targeted by drugs currently in pharmaceutical pipelines, as listed on the Pharmaprojects database, which has a total of 1,089 genes targeted by pipeline drugs. A total of 63 individual genes mapped to 52 different GWAS traits and drugs with the same or closely related indication to the GWAS traits (considered as matches). Conversely, 92 individual genes map to 51 GWAS traits and drugs with indications different from the GWAS traits (considered mismatches or potential drug-repositioning opportunities). Some genes are in both lists as they have multiple GWAS phenotypes that resulted in both a match to an existing indication and also a potentially novel indication.

referred to 1,099 gene names. Of these, 991 genes with recognizable HUGO gene names from Entrez Gene constituted the starting list for further analysis.

We next investigated how many of these genes were amenable to pharmacological modulation using small molecules (in other words 'druggable' by small molecules) or biopharmaceuticals (in other words 'biopharmable' using therapeutic antibodies or protein therapeutics) and compared these results with the entire genome. Out of 991 genes, 212 (21%) were considered druggable by small molecules, and 469 (47%) potentially biopharmable, defined here as being annotated with either a signal peptide or a transmembrane domain in ENSEMBL. These proportions are higher than those derived from the entire genome, which contains 3,191 potentially druggable genes (17%, $P < 5e^{-5}$) and 7,411 potentially biopharmable genes (38%, $P < 6e^{-9}$; **Fig. 1** and **Supplementary Methods**).

Next, we investigated if this excess in druggable or biopharmable genes among GWAS genes was explained by differences in the proportion of housekeeping genes. We obtained a set of housekeeping genes and observed that these genes are marginally underrepresented ($P < 0.10$) in the GWAS-selected genes (**Supplementary Methods**). Taken together, this analysis shows that GWAS genes are significantly more likely to be theoretically druggable or biopharmable targets than expected by chance. This observation prompted us to investigate which of the 991 GWAS genes are targeted by drugs already launched or in development (preclinical and clinical) and for what disease indication.

From the Pharmaprojects database (http://www.pharmaprojects.com/), a resource compiling worldwide drug pipeline data, we identified 1,089 genes (corresponding to 6% of the genome) being pursued as a target by a launched product, a candidate in clinical phase or in preclinical development (**Supplementary Methods** and **Supplementary Table 2**). This list of targets and associated drugs represent our pool to associate with genes linked to disease by GWAS. Terminated projects were not included in our analysis, though these could provide additional drug-repositioning opportunities in future analyses. Of the 991 selected GWAS genes, 155 (15.6%) had an associated drug project (**Fig. 1**). Compared with 1,089 of 19,258 human genes, the GWAS gene set is enriched 2.7-fold in targets pursued by drugs, which is more than expected by chance (15.6% versus 5.7%, $P < 3.5e^{-34}$; **Supplementary Methods**). This extends the theoretical expectation of GWAS druggability to practical application in drug discovery. This analysis included not only small molecules and biologics, but also other pharmacological modalities, such as antisense therapeutics. We then compared the disease indications pursued by the drugs with the GWAS trait for each of these 155 genes to identify matches and mismatches between the disease indications and the GWAS traits (**Fig. 1**). The analysis was done manually because the disease-related terms use different vocabularies in the catalog of GWAS data and Pharmaprojects (**Supplementary Methods**). We identified 97 matches between a drug indication and a GWAS trait corresponding to 63 individual genes and 52 GWAS traits; these observations could be considered as supportive for the particular indication being pursued. In addition, we also detected 123 mismatches, which included 92 individual genes and 51 GWAS traits (**Fig. 1**). These mismatches could be the basis for drug-repositioning opportunities. The entire lists are included in **Supplementary Tables 3** and **4**.

Table 1 contains six selected examples of matches. One of the most illustrative examples of an identical match is the 3-hydroxy-3-methylglutaryl-CoA reductase (*HMGCR*) gene. *HMGCR* is targeted by statins, a well-known class of cholesterol-lowering medications, and SNPs within this gene have been unambiguously associated with low-density lipopolysaccharide (LDL) cholesterol levels in GWAS data[1]. For less-advanced drug development programs, a match could provide more confidence in the disease indication. Examples include the following: (i) the human monoclonal antibody (mAb) ustekinumab (Malvern, Pennsylvania–based Centocor/Janssen-Cilag) targeting *IL12B*, currently marketed for psoriasis but only in phase 2 for Crohn's disease; or (ii) the preclinical program from Mellitech (Grenoble, Switzerland) for type 2 diabetes with small-molecule agonists targeting the solute carrier family 30 (zinc transporter) member 8 gene (*SLC30A8*).

Table 1 also includes examples where the GWAS trait is closely related, but not identical to, the disease indication for the drug of the same target gene. These imperfect matches may pinpoint the right disease indication for the drug. An example is the monoclonal antibody for the *ILR2A* gene (Novartis;

Table 1 Selected examples of matches between GWAS trait and drug indication[a]

Drug name or class	Most advanced development phase (for the indication)	Gene	Drug indication	GWAS trait	GWAS reference
Statins	Launched	HMGCR	Hypercholesterolemia	LDL Cholesterol	1
Ustekinumab	Approved	IL12B	Psoriasis	Psoriasis	13
Ustekinumab	Phase 2	IL12B	Crohn's disease	Crohn's disease	2
Anti-IL2 receptor mAb	Phase 2	IL2RA	Ulcerative colitis	Crohn's disease	2
AMG -785/ CDP-7851	Phase 2	SOST	Bone regeneration/ osteoporosis	Bone mineral density	14
Znt8 agonists	Preclinical	SLC30A8	Type 2 diabetes	Type 2 diabetes	15

[a]Examples are ranked from most advanced drug (launched) to less advanced (preclinical). The associated gene between each GWAS and the drug is shown. The drug indication and the phase of development for each drug are derived from the Pharmaprojects database. In each example the GWAS trait is identical (rows 1, 2, 3 and 6) or closely related (rows 4 and 5) to the drug indication. For the full list, see **Supplementary Table 3**. In many cases, more drugs for the gene are listed in the database at different phases. The GWAS references are from the catalog of GWAS data (http://www.genome.gov/gwasstudies).

Basel) in phase 2 to treat ulcerative colitis. Both Crohn's disease and ulcerative colitis are chronic inflammatory bowel diseases, but at the time of this analysis, *ILR2A* was only associated by GWAS with Crohn's[2] and not with ulcerative colitis. This suggests that pursuing that indication, in addition to ulcerative colitis, could be attractive.

Table 2 highlights six selected examples of drug-repositioning opportunities (that is, mismatches). For example, denosumab (Prolia, developed by Thousand Oaks, California–based Amgen/GlaxoSmithKline) is a marketed drug indicated for the treatment of postmenopausal women at high risk of fracture with osteoporosis. Denosumab targets the gene tumor necrosis factor (ligand) superfamily, member 11 (*TNFSF11*) also known as *RANKL*. *TNFSF11* has been associated with Crohn's disease by GWAS[2] and may potentially play a role in inflammatory bowel disease[3]. More work is required to understand mechanistically the role of *TNFSF11*, but it is still tempting to speculate that denosumab could be tested for Crohn's disease.

To investigate a potentially pleiotropic role of *TNFSF11* in Crohn's disease risk, we assessed whether the top allele (rs2062305[G]) associated with Crohn's disease[2] at *TNFSF11* was associated with differential expression of *TNFSF11* in human B-lymphoblastoid cells and human bone cells (osteoblasts) using direct assessment of *cis*-regulatory variation and *cis*-e quantitative trait locus analyses, respectively (**Supplementary Methods** and **Supplementary Table 5**). The Crohn's disease association at rs2062305 can explain population variation in *TNFSF11* expression in both cell types representing distinct cellular lineages relevant for both inflammatory (autoimmune) as well as bone disease. This provides additional evidence of a potential causal link between *TNFSF11* and Crohn's disease.

A phase 1 example is Biib-033 (Biogen Idec, Cambridge, MA, USA), which is an antibody targeting the leucine-rich repeat and immunoglobulin domain-containing 1 (*LINGO-1*) gene that is being developed for patients with multiple sclerosis. Two GWAS[4,5] studies suggest that *LINGO1* may also be an attractive target for essential tremor, a neurological disorder with limited treatment options[6].

The two previous examples have some information in addition to genetics to support the alternative indications. Other mismatches in **Table 2** have less additional support and require more investigation, but could provide surprising drug-repositioning opportunities. Such an example is for the dopamine beta-hydroxylase (*DBH*) gene and nepicastat (Roche; Basel), an inhibitor of *DBH* in phase 2 development for cocaine addiction and post-traumatic stress disorder. *DBH* has been associated by GWAS with smoking cessation[7]. It is thus tempting to speculate that DBH inhibitors may be beneficial for smoking cessation, while acknowledging that the direction of effect is not known yet.

It is important to stress that not all mismatches will lead to successful drug-repositioning opportunities. An illustration comes from the nitric oxide synthase 2, inducible gene (*NOS2*) with a range of *NOS2* inhibitors developed for various diseases, such as mucositis or rheumatoid arthritis. Recently, SNPs within *NOS2* have been associated in GWAS with psoriasis[8], raising the possibility that psoriasis may be an attractive indication, which is supported by the observations that skin lesions of psoriatic patients show an increase in nitric oxide production[9]. Contrary to these expectations, at least one small study did not identify any clinical improvement in psoriatic subjects when a topical inhibitor of nitric oxide synthesis was applied[10]. Negative studies are rarely definitive, but this example highlights some of the limitations of the approach proposed here.

GWAS signals in gene-rich loci, present another potential challenge, where linkage disequilibrium will often make it difficult to directly identify the causal gene, requiring individual scrutiny of each region. Our analyses are also fundamentally based on situations where the drug target matches a GWAS-identified locus. However, GWAS may affect the ligand, whereas drug discovery programs target the receptor (or vice versa). In addition, the direction of effect may differ between the GWAS gene and desired drug action. Additional work to build on GWAS

Table 2 Selected examples of mismatch between GWAS trait and drug indication[a]

Drug name	Most advanced development phase (for the indication)	Gene	Current drug indication	GWAS trait (new potential drug indication)	GWAS references
Denosumab/AMG-162	Launched/registered	TNFSF11	Osteoporosis/bone cancer	Crohn's disease	2
RPI-78M	Phase 3	IL27	Adrenoleukodystrophy	Crohn's disease/inflammatory bowel disease	2,16
Nepicastat	Phase 2	DBH	Cocaine addiction/post-traumatic stress disorder	Smoking cessation	7
Biib-033	Phase 1	LINGO-1	Multiple sclerosis	Essential tremor	4,5
AMG-557	Phase 1	ICOSLG	Systemic lupus erythematosus	Crohn's disease/celiac disease/ulcerative colitis	17–19
Cwp-231	Preclinical	TCF4	Cancer	Fuchs's corneal dystrophy	20

[a]Examples are ranked from most advanced drug (launched) to less advanced (preclinical). The associated gene between each GWAS and the drug is shown. The drug indication and the phase of development for each drug are derived from the Pharmaprojects database. For the full list, see **Supplementary Table 4**. In many cases, more drugs for the gene are listed in the database at different phases. The GWAS references are from the catalog of GWAS data (http://www.genome.gov/gwasstudies).

might include pathway analysis to better understand the biological system being manipulated by the drug, or indeed other forms of genetic analysis, such as Mendelian randomization, which has been successfully applied to establish the direction of causality for genes in diseases[11].

In summary, our methodology provides evidence that GWAS data do not only give insights into the biology of diseases, but may lead to immediate translational opportunities for drug discovery and development. Indeed, these genetics studies can pinpoint the right disease indication to targets and suggest concrete therapeutic opportunities for the repositioning of existing drugs in multiple therapeutic areas, including cancers, cardiovascular diseases, neurological disorders, immunological and inflammatory conditions and respiratory disorders. Our approach is limited to diseases investigated by GWAS. For example, for orphan diseases it would be more appropriate to identify targets using the Online Mendelian Inheritance in Man database that contains causal genes for Mendelian disorders or an alternative resource, such as the Rare Disease Repurposing Database[12]. It is also difficult to assess ascertainment bias in our study. The influence of GWAS on the initiation of clinical trials is challenging to determine because the rationale to start a drug discovery project is not easily available. It is possible that some of the GWAS have influenced clinical trials, leading to the 63 matches that would indicate an impact of genetics on drug discovery. Even so, the set of 63 matches cannot be entirely based on an ascertainment bias as some of these drugs were approved before the GWAS results were published (obvious examples include statins and glitazones). Because new studies are continuously added to the GWAS database and next-generation sequencing studies may reveal new gene-disease associations, additional drug-repositioning opportunities shall arise by careful monitoring of these developments.

Note: Supplementary information is available on the Nature Biotechnology website.

AUTHOR CONTRIBUTIONS
P.S., P.A. and M.R.B. conceived and designed the study, P.A. ran the computational pipeline, P.S., P.A., T.P., L.R.C. and V.M. analyzed the data, and P.S., P.A., T.P., J.B.R., M.R.B., L.R.C. and V.M. wrote the paper.

ACKNOWLEDGMENTS
The authors thank K. Peters for helpful comments on this manuscript. T.P. holds a Canada Research Chair and is supported by the Canadian Institutes of Health Research (CIHR). J.B.R. is supported by the CIHR, Ministere Développement Economique, Innovation et Exportation du Québec, Fonds de la Recherche en Santé du Québec, Lady Davis Institute of Medical Research.

COMPETING FINANCIAL INTERESTS
The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturebiotechnology/.

Philippe Sanseau[1], Pankaj Agarwal[1], Michael R Barnes[1,5], Tomi Pastinen[2], J Brent Richards[3], Lon R Cardon[4] & Vincent Mooser[4,5]

[1]Computational Biology Department, Quantitative Sciences, GlaxoSmithKline, Stevenage, UK. [2]Departments of Human and Medical Genetics, McGill University and Genome Quebec Innovation Centre, Montreal, Quebec, Canada. [3]Departments of Medicine, Human Genetics, Epidemiology and Biostatistics, Lady Davis Institute, Jewish General Hospital, McGill University, Montreal, Quebec, Canada. [4]Genetics Department, Quantitative Sciences, GlaxoSmithKline, King of Prussia, Pennsylvania, USA. [5]Present addresses: William Harvey Research Institute, Queen Mary University of London, Charterhouse Square, London, UK (M.R.B.) and Department of Pathology and Laboratory Medicine, CHUV University Hospital, Lausanne, Switzerland (V.M.).
e-mail: philippe.x.sanseau@gsk.com

1. Kathiresan, S. *et al. Nat. Genet.* **40**, 189–197 (2008).
2. Franke, A. *et al. Nat. Genet.* **42**, 1118–1125 (2010).
3. Moschen, A.R. *et al. Gut* **54**, 479–487 (2005).
4. Stefansson, H. *et al. Nat. Genet.* **41**, 277–279 (2009).
5. Clark, L.N. *et al. Eur. J. Hum. Genet.* **18**, 838–843 (2010).
6. Louis, E.D. *Lancet Neurol.* **4**, 100–110 (2005).
7. Tobacco and Genetics Consortium. *Nat. Genet.* **42**, 441–447 (2010).
8. Stuart, P.E. *et al. Nat. Genet.* **42**, 1000–1004 (2010).
9. Ormerod, A.D. *et al. Arch. Dermatol. Res.* **290**, 3–8 (1998).
10. Ormerod, A.D. *et al. Br. J. Dermatol.* **142**, 985–990 (2000).
11. Welsh, P. *et al. J. Clin. Endocrinol. Metab.* **95**, 93–99 (2010).
12. Xu, K. & Coté, T.R. *Brief. Bioinform.* **12**, 341–345 (2011).
13. Nair, R.P. *et al. Nat. Genet.* **41**, 199–204 (2009).
14. Styrkarsdottir, U. *et al. Nat. Genet.* **41**, 15–17 (2009).
15. Zeggini, E. *et al. Science* **316**, 1336–1341 (2007).
16. Imielinski, M. *et al. Nat. Genet.* **41**, 1335–1340 (2009).
17. Barrett, J.C. *et al. Nat. Genet.* **40**, 955–962 (2008).
18. Anderson, C.A. *et al. Nat. Genet.* **43**, 246–252 (2011).
19. Dubois, P.C. *et al. Nat. Genet.* **42**, 295–302 (2010).
20. Baratz, K.H. *et al. N. Engl. J. Med.* **363**, 1016–1024 (2010).

Bias in high-tier medical journals concerning physician-academic relationships with industry

To the Editor:

Economists have concluded that extensive interaction between academic researchers and practicing physicians with industry has facilitated the development and dissemination of biomedical diagnostics, devices and therapies[1–3]. However, academic health centers, states and the federal government have instituted regulations designed to monitor, limit or eliminate such interactions based on concerns that a downside is that such relationships may degrade the performance and reporting of biomedical research and also induce physicians to behave in a manner inconsistent with cost-effective or ethical patient care—which are loosely defined under the operational term 'financial conflicts of interest' (COIs).

As general medical journals are major repositories of medical policy discussion, we sought not only to analyze whether the positive and negative aspects of industry-academic relationships were equally represented in top-tier medical journals but also to assess the weight of evidence in the COI literature that patient outcomes or public attitudes are indeed negatively affected by corporate interactions with academics and physicians.

We analyzed papers published in four journals selected on the basis of their high (>20) impact factors: The *Journal of the American Medical Association* (*JAMA*), *The Lancet*, *Lancet Neurology* and The *New England Journal of Medicine* (*NEJM*). We identified articles by performing keyword searches on the websites of the selected journals and on PubMed (US National Library of Medicine, National Institutes of Health). We obtained additional articles from references cited in the papers culled from the keyword searches. Keywords included 'conflict of interest', 'CME' (continuing medical education), 'physician-industry relationships', and 'academic medical center (AMC)-industry relationships'.

The 108 articles selected for analysis (**Supplementary Data**) encompassed reports