Crowdsourcing in pharma: a strategic framework

Jörg Bentzien[1], Ragu Bharadwaj[2] and David C. Thompson[1]

[1] Boehringer Ingelheim Pharmaceuticals, 900 Ridgebury Road, Ridgefield, CT 06877, USA
[2] Appirio, 760 Market St, 11th Floor, San Francisco, CA 94102, USA

Conceptually, all organizations can be described as coordinated actors working together to deliver a product(s), or provide a service(s). For organizations to remain competitive, it is important to have processes that look outward for external 'innovations' that could improve how work is done, and what is delivered. We present a comprehensive review of a variety of processes that pharmaceutical companies have used to engage external actors ('the crowd') to provide innovation in the service of delivering novel therapeutic agents. This culminates in a framework that provides a consolidated view of crowdsourcing processes, which in turn enables a strategic application of a crowdsourcing methodology based on problem type.

Introduction

The need to innovate

The role of pharmaceutical companies within the healthcare ecosystem is in the provision of safe and efficacious treatments that positively affect patient quality-of-life. The discovery and development of these treatments is a complex, time-intensive and costly endeavor often running into billions of dollars over 10–15 year cycle times and with a very low rate of success. Given its high costs, capital markets have proven very useful in funding the majority of pharmaceutical companies. Unfortunately, such capital is neither patient nor long-term. This has further complicated the roles of pharmaceutical companies, which also have to satisfy shareholder demands for capital appreciation, certainty and quick returns. Despite significant advances in the science of R&D, along with commensurate improvements in technological and managerial factors, all things that should enable increased efficiency in commercial drug R&D, the number of new drugs approved per billion dollars spent has halved roughly every nine years since 1950 [1]. Furthermore, the rate of approvals is below that required to generate sufficient growth for the industry as a whole [2].

There are many reasons for this, and there is much discussion in the literature as to what ails the pharmaceutical sector and myriad ways suggested to potentially fix it [3]. It is important to remember however that the search for drugs is one that is occurring in an unfathomably large search space – estimates range between 10^{23} and 10^{60} potentially realistic drug-like molecules that are synthesizable [4]. Finding a novel, commercially viable product that exhibits superior efficacy and safety compared with existing treatment options ensures that pharmaceutical endeavors remain extraordinarily risky ventures unfolding in a context of incomplete knowledge.

A common strategy to manage risk is the adoption of a portfolio approach, wherein a basket of known quantities sits aside more experimental approaches. Given the lengthy timescales involved in the drug discovery process, the importance of a robust assessment of the target is crucial [5]; the more information on the target and its viability, the more of a 'sure thing' one might suppose it to be. Such known quantities are likely to experience steep competition as the market arranges itself accordingly, so differentiation at the portfolio level is vital.

Risk-focused portfolio management is an example of the underlying tension between exploitation and exploration – known quantities and novelty – and represents the most important challenge any organization has to wrestle with throughout the span of its existence. Is the balance of exploitation (using what is already institutionally known and accepted) and exploration

Corresponding author: Thompson, D.C. (david.thompson@boehringer-ingelheim.com)

http://dx.doi.org/10.1016/j.drudis.2015.01.011

(investigating what is not institutionally known and accepted) appropriate? [6,7]. Another interesting framing of this same dichotomy is internal versus external. How much of what is done within an organization is planned, sourced and executed internally – versus similar activities with a focus more external to the organization.

Outside in

The modern global pharmaceutical industry has its 19th century origin in two sources: apothecaries that transitioned into the wholesale manufacture of drugs and chemical companies that established research-oriented laboratories focused on the medical application of their products. Cooperative relationships between academic laboratories and pharmaceutical firms were established early on, and drove a focus on dyes, antibodies and physiologically active agents [8]. These relationships are examples of the first instance of processes that internalize external innovation. If one broadly defines 'crowd' as the agents external to an organization, this is arguably the first use of crowdsourcing in the pharmaceutical space.

Crowdsourcing

Crowdsourcing is a term that was coined in a 2006 issue of *Wired* magazine [9] and described an internet-enabled business model that harnessed the creative ability of agents external to an organization. As implied above, crowdsourcing existed before the internet and one of the best-known examples of crowdsourcing, pre-internet, was the British government's establishment of the Longitude Act in the 18th century. To prevent the loss of ships at sea, the government created a prize purse of £20 000 (the equivalent of £2.5 million in 2014) to map longitude. The winning solution, the chronometer, came from an unexpected source, John Harrison, a carpenter and clockmaker by trade, and was delivered some 50 years after the establishment of the act [10]. The unexpected nature of the winning solution is a result of using a process that enables exploration and demonstrates that, when constructed appropriately, such searches encourage but do not necessarily reward 'expert bias' [11]. The use of crowdsourcing has grown following the widespread adoption of the internet. The ready access to a distributed network has driven the widespread exposure of problems and the identification of solvers.

In this present work, we extend the definition of 'crowd' to include any actors external to an organization, working with or for the organization and in the service of solving problems of interest to the organization. In doing this we are able to connect a variety of processes that internalize external innovation (and that hark back to the origins of the pharmaceutical industry) into a comprehensive framework. Such a framing coherently connects open innovation, crowdsourcing, academic collaboration(s), consortia and pre-competitive participation activities into a single vision, amenable to strategic use.

This paper is organized accordingly. In the following section crowdsourcing is described in more detail, along with recent examples of the application of crowdsourcing to problems in informational R&D (inside and outside of the pharmaceutical industry). Some thoughts are presented on the role of community and the importance of domain abstraction, along with a brief discussion of when crowdsourcing might not work. Following this, a framework is introduced that rationalizes engagement of the crowd through crowdsourcing with other processes that have previously been employed by the pharmaceutical industry in an effort to ensure efficient internalization of externally (to the organization) innovative practices. Concluding remarks are then offered.

Crowdsourcing examples

Following the definition of crowdsourcing presented above, in Fig. 1 we outline a variety of ways through which companies are currently crowdsourcing work via internet-enabled services. We have decomposed this along a 'complexity of task' axis, ranging from micro-tasks that can be performed in seconds using a service like Amazon's Mechanical Turk (https://www.mturk.com/mturk/welcome) through to multi-hour research activities on platforms such as InnoCentive® (http://www.innocentive.com/), Kaggle™ (http://www.kaggle.com/) and [topcoder]™ (http://www.topcoder.com/). It is important to note that the focus below is on 'informational' crowdsourcing, wherein the crowd is solely engaged with information and the activity and their participation is solely digital. By contrast, there are a variety of recent examples of crowdsourcing with a material physical component, wherein the material is sourced through the crowd: examples include soil (mySoil) [12], fecal matter (The American Gut Project; http://humanfoodproject.com/americangut/) and genetic material (The Resilience Project; http://resilienceproject.me/).

A second dimension in Fig. 1 describes the level of subjectivity applicable to a particular crowdsourced solution. A crowdsourcing platform that enables the objective improvement and optimization of a codebase presents a different (albeit no less useful) resource than a service optimizing the inherently subjective effectiveness of a marketing campaign. In general, with objective metrics to define results, crowdsourcing becomes more amenable to contests through an association of points and a subsequent ranking. When the level of subjectivity is high, collaborative mechanisms and nonmonetary rewards such as pro-social membership seem to dominate. In addition to these dimensions, a variety of additional factors (such as problem type, task modularity, task virtualization and the ability to attain a competitive advantage) underlie the motivations and explain the logic behind when to use crowdsourcing and how.

Across industries and platforms crowdsourcing consists of several common elements: a well-defined contest statement describing the context of the problem is shared. The statement includes details regarding the size of any prize, the method of evaluation and the duration of the activity. The problem statement, the method of evaluation and all additional considerations discussed briefly in the prior paragraph determine where in Fig. 1 a crowdsourcing activity will lie. We detail below some current crowdsourcing examples and successes across fields, and then focus on the pharmaceutical industry. These examples highlight the use of crowdsourcing for finding innovative solutions to problems in R&D; many examples exist of crowdsourcing being used outside the R&D sector [13,14].

The US National Aeronautics and Space Administration (NASA) has consistently used crowdsourcing to solve hard innovation problems and to develop complex software solutions. Recent successes include algorithms to help optimally position the

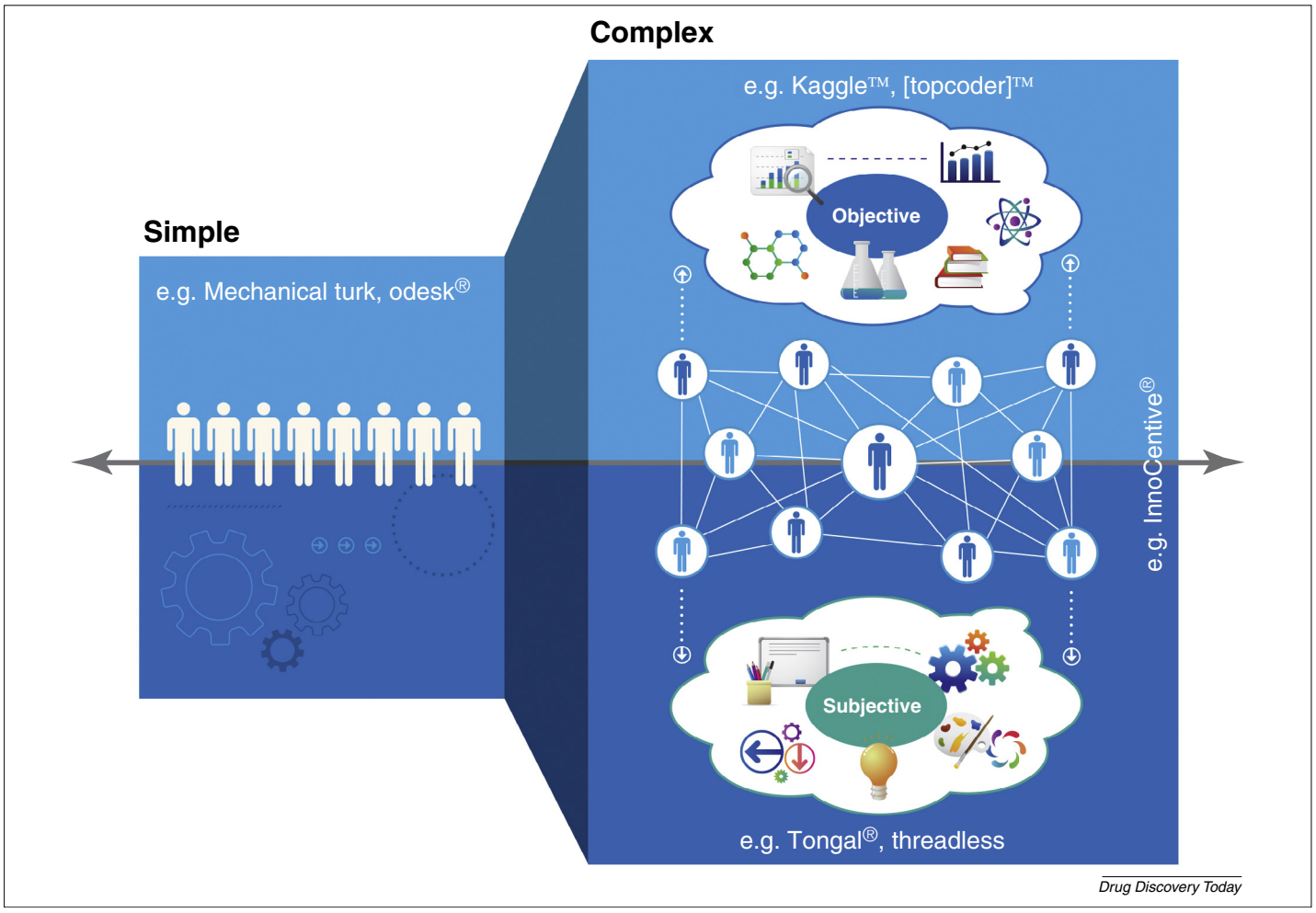

FIGURE 1

Informational crowdsourcing decomposed along a horizontal 'complexity of task' axis, ranging from micro-tasks that can be performed in seconds through to multi-hour research activities on platforms. A second, vertical axis arranges crowdsourcing platforms according to their 'flavor' – are problems hosted on these services amenable to 'objective' or 'subjective' optimization?

International Space Station (ISS) solar panels to increase the power generated while still working to eliminate shadowing effects [15], and in the development of algorithms to detect and track asteroids based on satellite images [16]. The US Environmental Protection Agency (EPA) has also used crowdsourcing to develop predictive analytics solutions for problems involving toxicity prediction for compounds [17,18] and detecting algal blooms within ponds and lakes using images [19]. Over the past two years General Electric, using a program called FlightQuest, has regularly engaged with multitudes of programmers and data analysts to develop algorithms to optimize flight plans and to predict flight arrival times efficiently [20].

Pharmaceutical companies have also used crowdsourcing solutions to solve data-rich problems, and there have been a number of applications across the research, development and clinical value chain. Some recent examples from the literature include the development of predictive cytotoxicity models [11], algorithmic improvements to a popular open-source genome-wide association study (GWAS) approach [21–23] and the application of deep learning to compound selection [24]. Other examples include the prediction of exacerbations in patients with respiratory disease [25], the identification of novel treatment approaches to cure inflammatory bowel disease [26], the development of biomedical devices [27] and in the identification of cheaper synthetic routes [28].

The importance of community and domain abstraction

Crowdsourcing platforms harness the collective cognitive surplus of solvers looking to bring their time, skills and energy to the exposed problems [29]. With respect to crowdsourced competitions, the community affords a context within which such competitions unfold. Individuals compete against each other, perhaps receiving feedback on their progress, until the end of a competition. Participation can be aggregated into a history and even a score, which in turn affords unique opportunities within a community or becomes a transferable badge of accomplishment outside a community [30].

Such a community provides a resource that can be engaged by an organization, and often communities have a focus that can direct an organization's attention based upon the problem at hand, for example software ([topcoder]™), data science (Kaggle™, [topcoder]™, Tunedit; http://tunedit.org/), computational biology (DREAM; http://dreamchallenges.org/) or creative work (Tongal®; https://tongal.com/).

One particularly successful approach to engagement when looking for broad participation of solvers is domain abstraction. In such an approach the work is abstracted from its originating domain and is presented as a problem in another. This relies on underlying isomorphic characteristics of the problem. Such an approach either broadens participation, increases participation or both. A particularly interesting abstraction is that explored by the fold.it (http://fold.it/portal/) platform that has translated the problem of protein folding and design into a highly graphical endeavor that a large number of people can 'play', without explicit knowledge of modeling or protein structure and/or function. This application of domain abstraction has resulted in the solution of several outstanding problems of scientific interest as well as algorithmic improvements [31,32].

In another example of abstraction, the problem of the creation of predictive models of cytotoxicity was removed from the computational chemistry community and re-framed as a data science challenge, free of chemical descriptors and identifiers and presented as an exercise in machine learning [11]. The phasing problem in crystallography has been similarly abstracted and presented in a graphical and engaging fashion [33].

Domain abstraction combined with the inclusion of competitive elements in the context of community, work together to enrich the likelihood that a near optimal solution, for any given community of solvers, is discovered. These are powerful, albeit empirical, proofs of the 'diversity trumps ability' theorem conjectured (and formally proven) by Page and Hong [34].

In a final note on this topic, abstraction has recently been taken to the next level by connecting online and offline activities and allowing players of the EteRNA 'game' to remotely carry out real experiments to verify computational predictions of how RNA molecules fold. Here, not only has the problem of RNA folding been abstracted to an engaging game open to non-subject matter experts, but the digital insight has been translated back into the real world [35,36]. This occurs through the physical synthesis of molecules predicted, computationally, to be stable and connects the material and informational components of crowdsourcing described above.

When crowdsourcing does not work

Although the above examples are compelling, it is as important to recognize when one cannot employ the power of the crowd:

- If the solution holders are accessible to you. In such cases you are probably working with the holders of the information required to generate a solution, so engaging a crowd is unproductive.
- When the availability of a crowd of competent solvers is limited.
- If you do not have an internal framework established to evaluate the solution(s).
- If the suspected cost (money and/or time) of implementation of the solution(s) is too high (although this is mitigated by what is learnt through engaging the crowd, as experienced by Netflix) [37].
- If safeguards are not in place to prevent malicious or ill-aligned participation from the crowd (e.g. the submission of erroneous 'red balloon' locations during the DARPA challenge) [38].
- If the solution(s) are embedded in processes containing lengthy feedback loops.
- In the instance where intellectual property (IP) rights are of importance, special mutual agreements need to be considered and enforced. This is of particular importance to the pharmaceutical industry, as reliant as it is on IP.
- When the information required to solve a problem might not be present in the crowd or completely known. This can happen with, for example, predicting the success of drugs in clinical trials or in predicting industry standards that will attain widespread adoption (e.g. the USB standard) [39].

Strategic framework

We have described many of the characteristics of internet enabled crowdsourcing, with a focus on R&D and the pharmaceutical industry. This represents one example of how organizations are able to internalize external innovation. What follows, for the remainder of this contribution, is a description of a strategic framework that aligns crowdsourcing with other, more traditional, elements of how the pharmaceutical industry looks outside itself.

The principal element of this framework is the 'organization'. We assume that this is a collection of individual actors, engaged in 'work' that ultimately results in the delivery on its promise to its customers in the provision of a service(s) or product(s). For this work, actors not formally part of the organization, with no role in work directly relating to the delivery of the service(s) or product(s) consumed by customers, are to be considered as 'other'.

Accordingly, one can arrange relationship patterns between organization and other according to an abstract coupling constant. Such a coupling constant, represents how closely connected organization is to other, and arranges recent crowdsourcing activities with more-traditional approaches to insourcing external innovation. One key factor that mediates the strength of the connection between other and organization is the overlap between other and organization objectives and incentives. This is further illustrated in Figs. 2–4 and in the following sections, wherein each regime of interaction is described in more detail.

Strongly coupled

An example of a strongly coupled relationship between organization and other is a hosted postdoctoral fellowship program (Fig. 2). Such programs have been around for a long time, represent a bridging pathway between academia and industry and are ubiquitous throughout the pharmaceutical industry [40]. Such programs can have an integral role in the talent management strategy of an organization, providing useful vocational experience. Usually with a basic science focus, the postdoctoral associate is tasked with publishing novel research and is typically not connected to traditional drug discovery efforts. Reviews of the top five pharmaceutical company postdoctoral offerings (by 2013 global sales) highlight this focus on internalizing external innovation through the use of these programs:

- 'Our program provides postdoctoral scholars with a unique opportunity to perform innovative fundamental research in a pharmaceutical setting' (http://postdoc.nibr.com/).
- 'Trainees pursue their research and career training in a culture steeped in the translation of basic science discovery into difference-making medicines for patients of all ages and geographies.' (http://pfizercareers.com/university-relations/postdoc).

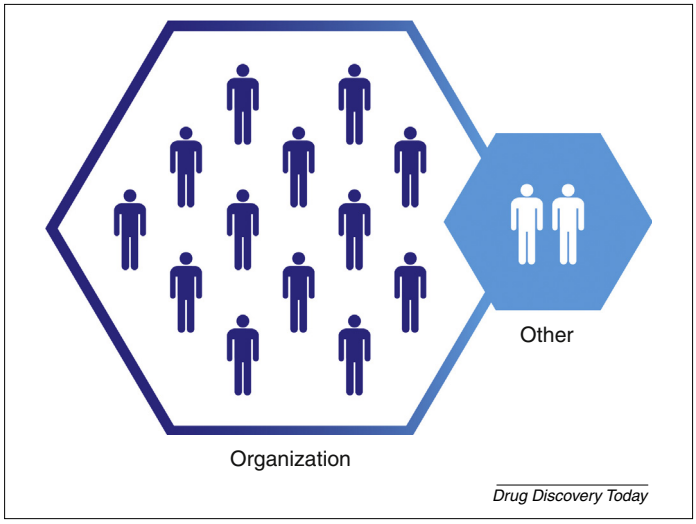

FIGURE 2

In the strongly coupled regime, the overlap between 'other' and 'organization' objectives and incentives is strongly aligned. Examples include hosted postdoctoral fellowships. Such programs have been around for a long time, represent a bridging pathway between academia and industry and are ubiquitous throughout the pharmaceutical industry [40].

- 'Generate innovative science resulting in high-quality external publications' (http://www.merck.com/research/fellow/home.html).

As such, they can be considered other in the definition above, given that their primary role is removed from the ultimate provision of existing service(s) or product(s).

Accordingly, postdoctoral programs can be considered an abstraction from an academic setting into an organizational one. The supposed expectation is that associates will bring with them cutting-edge perspectives on community best practices and innovations and a vibrant and active network of collaborators. To our knowledge there is no literature on the efficacy of these programs from an internalization of innovation perspective. In this strongly coupled regime, the overlap between other and organization objectives and incentives is strongly aligned.

Partially coupled

Partial coupling refers to the situation where other is connected to organization, but there can only be partial overlap between mutual objectives and incentives. For the purposes of the current framework we explore two different examples of this regime: (i) consortia and (ii) academic collaboration and open innovation.

Consortia

In this example, actors from within an organization, with appropriate subject matter expertise, are abstracted from the organization, along with other, similar actors from other organizations, and enter into a consortium (Fig. 3a). In this context, they interact with others, external to their own organization in the service of something larger than any of the contributing participants. Illustrative non-pharma examples would include IBM's partnership with the Linux and the Apache foundations [14]. In both instances, IBM capital (human and financial) was deployed in the development of the Linux and Apache communities; something of use to the organization (IBM) but 'bigger' than it.

There are numerous examples in the pharmaceutical space. Some of these have a very broad scope and are government-funded like the European Innovative Medicines Initiative (IMI) [41], which includes multiple programs like the Joint European Compound Library [42] and Open PHACTS [43], others are more focused like the Structural Genomics Consortium (http://www.thesgc.org/) [44], or geared toward specific diseases [45]. Yet, other consortia are concerned with setting industry standards (e.g. the Controlled Substance Compliance Program headed by the Pistoia Alliance) [46].

The IMI was launched in 2008 as a public–private partnership between the European Commission and the European Federation of Pharmaceutical Industries and Associations (EFPIA). With a total budget of €2 billion, the IMI supports 46 collaborative research projects with the aim of developing new technologies and methodologies for pharmaceutical R&D. These projects bring together experts from industry, academia, small and medium enterprises (SMEs), regulators and patient groups [47,48].

The European Lead Factory (ELF), which was launched in January 2013, is one of the projects supported by the IMI. This collaboration brings together 30 academic and industry partners including Bayer, Janssen, Merck-Serono, AstraZeneca, Sanofi, UCB and Lundbeck. The European Lead Factory was established to promote the discovery of novel small molecule candidates, suitable for subsequent optimization either to drug candidates or to high-quality pharmacological tools for the experimental validation of targets. To facilitate this goal a chemical screening library was established. This library consists of 300 000 chemical compounds donated from the seven EFPIA partners. In addition, the ELF has the goal to add another 200 000 compounds through collaborative efforts from academia and SMEs. Both libraries together form the Joint European Compound Library (JECL). Academics as well as SMEs and patient organizations can propose targets to the ELF for screening against the JECL. Accepted targets are screened for free and a qualified hit list of up to 50 compounds is returned to the program owner. The program owner gains full access rights and three-year exclusivity to exploit the screening results. In the case of direct exploitation, not for research use, the program owner will compensate the ELF via a scheme of milestone payments [49].

Such consortia play an interesting role in fostering collaboration and in pooling expertise across organizations [50]. Although, when viewed as an abstraction from organizations, perhaps unsurprisingly, the issue of 'consortium fatigue' [51] surfaces. Put simply, there is no agency tasked with coordinating consortia in a way that ensures optimal collaboration or use of resources across all consortia. This can lead to confusion and inefficiency. As is remarked succinctly in [51], 'we need an evidence-based approach – a science of collaboration – to evaluate and inform the evolving multi-stakeholder collaboration environment in biomedical innovation'.

Academic collaboration and open innovation

As was remarked in the introduction to this work, academic collaboration has been a component of the pharmaceutical industry since its inception. Connections between industry and academic organizations are well studied [52,53] and there typically exists a spectrum of one-off relationships with external collaborators, through to an engagement of open innovation programs. This is illustrated in Fig. 3b.

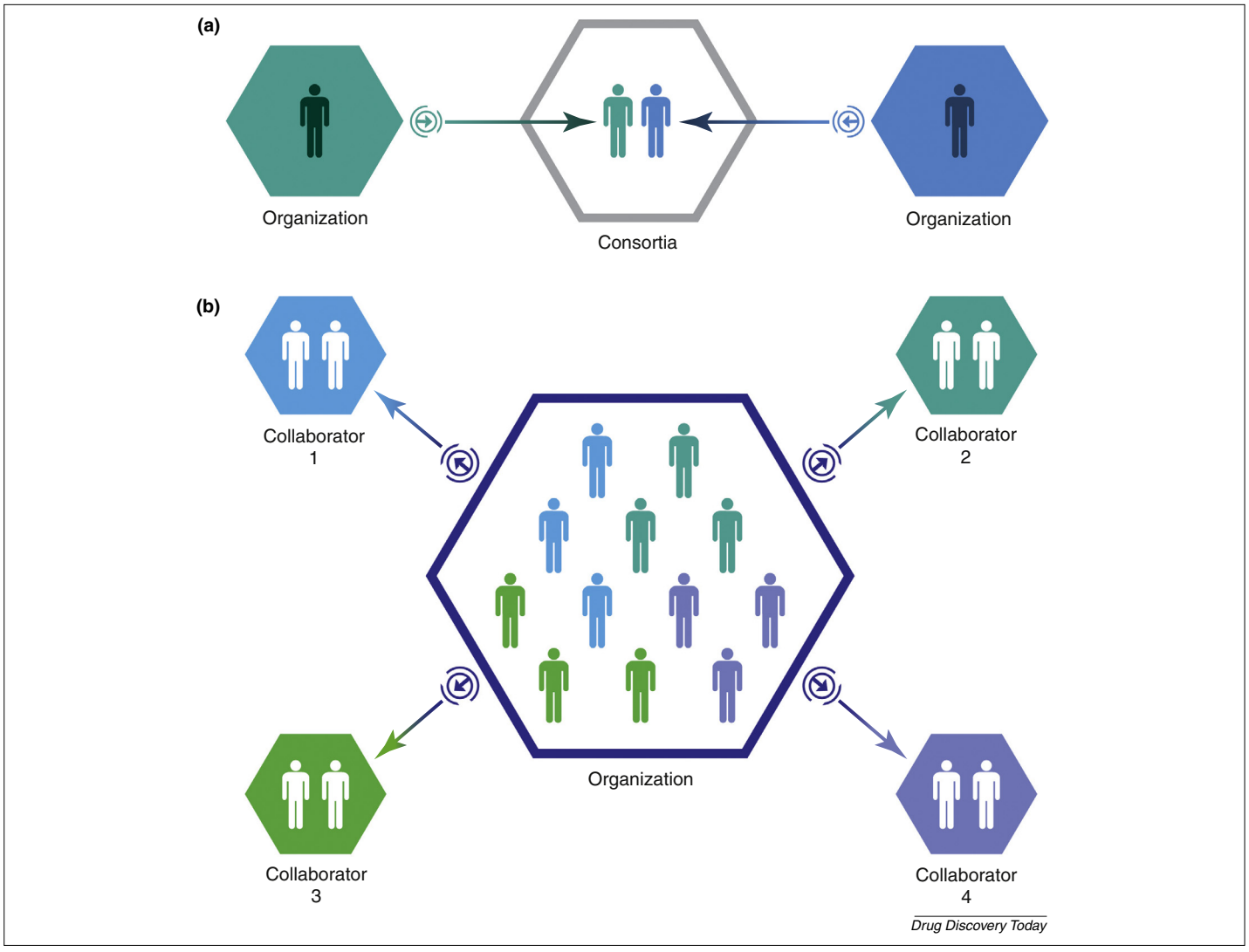

FIGURE 3

Partially coupled regimes. **(a)** Actors from within an 'organization', with appropriate subject matter expertise, are abstracted from the organization, along with other, similar actors from other organizations, and enter into a consortium. In this context, they interact with 'others', external to their own organization in the service of something larger than any of the contributing organizations. A variety of examples in the pharmaceutical space can be found in [41]. **(b)** A network of external-facing collaborations between an organization and a variety of, for example, academic investigators. There is an opportunity to coordinate 'point engagements' of academic collaborators because they could represent a nascent network partially coupled to the organization.

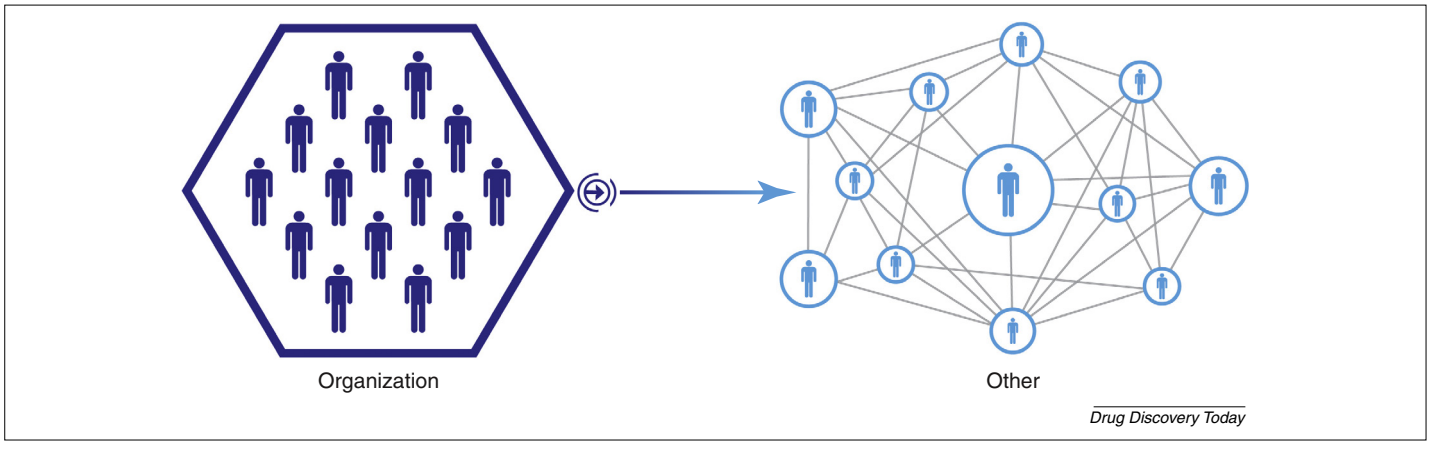

FIGURE 4

Weakly coupled regime. Other in this instance refers to the crowd: a loosely connected network of individuals, who, through internet-enabled services and community, can be engaged in a full spectrum of activities ranging from micro-task through to multi-hour research project.

TABLE 1

An overview of the different approaches discussed above, together with selected examples. To reiterate, the framework orients processes that seek to internalize external innovation, with respect to each other through use of an abstract coupling constant representing the 'strength' of interaction between 'organization' and 'other'.

Weakly coupled: crowdsourcing platforms

Name	Type of challenges	Audience of solvers	Examples and Refs	url
Kaggle™	Machine learning data challenges	Data scientists	Merck: molecular activity challenge [24] Boehringer Ingelheim predicting a biological response [11]	http://www.kaggle.com/
Sage Bionetworks	Bioinformatics data challenges	Bioformaticians, data scientists	Sage Bionetworks – DREAM breast cancer prognosis challenge (BCC) [68]	http://sagebase.org/
InnoCentive®	Diverse challenges Data challenges and reduction to practice	Diverse community of solvers	AstraZeneca challenge: novel biomarkers for neuropathic pain (ideation) Seeking substances with activity on the nicotinic acetylcholine receptor (reduction to practice) Prize4Life [69]	http://www.innocentive.com/
[topcoder]™	Coding and data challenges	Data scientists and programmers	EPA ToxCast™ prediction challenge [18]	http://www.topcoder.com/

Partially coupled: academic collaboration and open innovation

Name	Company	Audience	Example or Refs	url
Grants4Targets™	Bayer Healthcare	Academic groups and start-ups	[70,71]	http://www.grants4targets.com/scripts/pages/en/index.php
Open Innovation Drug Discovery (OIDD)	Eli Lilly	Academic groups and start-ups	[72]	https://openinnovation.lilly.com/dd/
Centers for Therapeutic Innovation (CTI)	Pfizer	Academic groups		http://www.pfizer.com/research/rd_partnering/centers_for_therapeutic_innovation
Discovery Fast Track	GSK	Academic groups		http://openinnovation.gsk.com/

Partially coupled: consortia

Name	Purpose	Example or Refs	url
Structural Genomics (SGC)	A public–private partnership that supports the discovery of new medicines through open access research	In coordination with GSK, identified the potential of Brd4 as a drug discovery target [44]	http://www.thesgc.org/
Innovative Medicines Initiative (IMI)	European public–private initiative aiming to speed up the development of better and safer medicines for patients	European Lead Factory Joint European Compound Library [42,47–49]	http://www.imi.europa.eu/
Open PHACTS	Discovery platform to reduce barriers to drug discovery in industry, academia and small businesses	[43]	http://www.openphacts.org/
Pistoia alliance	Lowering barriers to R&D innovation	Controlled substance compliance project [46]	http://www.pistoiaalliance.org/

Within an organization there exists a responsibility to coordinate 'point engagements' of academic collaborators because, on some level, they themselves could be considered a nascent network coupled to the organization. Consider, for example, all of the external collaborators working in the fields of fluidics or diabetes. These collaborators should be considered a resource, bound through joint scholarship in a field, and an internal network perspective on this resource could then become a strategic differentiator. At this time, we are not aware of research evaluating the efficacy of such partially coupled activities.

The pharmaceutical industry has embraced a wide variety of ongoing open innovation activities. Open innovation efforts look to engage, evaluate and align with the academic community in an effort to internalize external innovation. Our goal here is not to reiterate what has been perfectly well described previously but to connect open innovation as an example of an engagement with other that can be reconciled with alternate approaches that internalize external innovation. Comprehensive descriptions and discussions on this topic can be found in [2].

To tap into the entrepreneurial zeitgeist, and to accommodate the need for a material evaluation of the explorative activity (something that is markedly different from entrepreneurship in the high technology – software space), a variety of incubators have been established and staffed, for example Janssen Labs (http://www.janssenlabs.com/) or the 'outcubator' model of BioMed X [54]. When examined in this fashion, open innovation activities are also an example of partially coupled engagements with other(s).

Weakly coupled

Much of the introductory material in this work covers the weakly coupled regime of the proposed framework, which is illustrated in Fig. 4. As has been described, other in this instance refers to the crowd: a loosely connected network of individuals who, through internet-enabled services and community, can be engaged in a full spectrum of activities ranging from micro-task through to multi-hour research project (Fig. 1). As was described, such actors could associate with skills-based communities that have no connection with the pharmaceutical industry or healthcare.

An overview of the different approaches discussed above, together with selected examples, is summarized in Table 1. To reiterate, the framework orients processes that seek to internalize external innovation, with respect to each other, through use of an abstract coupling constant representing the strength of interaction between organization and other.

Concluding remarks and discussion

In this contribution the current application of internet-enabled crowdsourcing platforms to informational problems in R&D, inside and outside of the pharmaceutical industry, has been reviewed. Subsequently, this capability has been rationalized, through a simple framework, with other practices that have been employed to internalize external innovation. Such a framework is oriented with respect to a coupling constant that describes some measure of interaction between organization and other. This coupling constant is a proxy for a host of dimensions (some of which will be competing and conflicting), nonetheless providing an intuitive arrangement of innovation activities and processes available to organizations.

The current state of the R&D-based pharmaceutical industry has developed business models that are highly reliant on the capture of intellectual property (IP) and exclusivity as a means for generating funding across the process. This is especially important given the high costs of the regulatory process and clinical trials. Although the industry has clearly shown a proclivity toward open innovation and collaboration, its business models have not evolved rapidly enough to keep up. The framework as presented is rooted in the current state of the pharmaceutical industry – an IP-oriented for-profit endeavor. What about alternate models of drug discovery?

Recently, many have begun to argue that the traditional drug discovery model is ripe for reform [55]. The current business model of the pharmaceutical industry is unable to meet all medical needs, especially in the areas of neglected and rare diseases. In response, a variety of open source models have been implemented by a variety of organizations [56–58] (http://opensourcemalaria.org/; http://www.osdd.net/) and there has been some recent discussion regarding the viability of such an open source business model expansion beyond the area of neglected diseases [59,60] (http://www.opensourcepharma.net/). An interesting approach for drug development is the crowdsourcing of clinical trial design [61,62].

An open source model is predicated on the availability of data and the community development of open source tools. In the context of our framework, such an open source entity would operate solely through partially or weakly coupled activities, wherein the participation of other would be a vital driving force, although each actor can never devote their full attention to any one of the institution's problems or issues. In this regime, and operating under an open source approach, the role of incentives becomes increasingly important. In the for-profit example, all partially and weakly coupled activities are driven (in an energetic sense) by the organization. To ensure the sustainability of any open source approach, such considerations are vital – in particular how to fund and generate incentives in the absence of IP. This could be done through financing by government or private not-for-profit organizations (e.g. WHO, Bill and Melinda Gates Foundation, Wellcome Trust) [63], public–private partnerships, through crowdfunding [64] or pro-bono work by individuals. As discussed earlier, pro-social motivation can be incited through public recognition of individual contributions.

Assuming funding questions can be solved and individual steps in a drug discovery process can be executed through this open source model, there remain significant challenges. Drug discovery requires close interaction of people with diverse skill sets and can only succeed with fast feedback loops between material experimentation. Thus, an open source pharma model, using processes in the partially or weakly coupled regime, faces significant organizational challenges in coordinating the different elements while keeping the individual participants motivated and focused.

Whether through the traditional or nontraditional models, one might expect an increased inclusion of other in some 'future of work'. This has the potential to disrupt the employer–employee relationship and/or expectation significantly, leading to much looser, more-flexible voluntary relationships [65]. A wonderful exploration of one possible destination for this journey, specific to the life sciences, can be found in [66].

The scale and complexity of the discovery and development of novel, efficacious, tolerable and viable pharmaceutical products demands that we critically and continually evaluate how we continue to bring 'outside in'. This is a difficult problem for individuals [67], let alone organizations, but it remains vital. It is our hope that the suggested framework is used as an organizing principle by those looking to rationalize novel crowdsourcing methodologies with existing approaches that are deeply entrenched in how the pharmaceutical industry has developed, and will continue to develop in the future.

Acknowledgments

D.C.T. would like to thank Michael J. Montgomery and NYC Bio for the invitation to speak about crowdsourcing in drug development. It was during the preparation for this event that much of this framework was determined.

References

1 Scannell, J.W. *et al.* (2012) Diagnosing the decline in pharmaceutical R&D efficiency. *Nat. Rev. Drug Discov.* 11, 191–200
2 Schuhmacher, A. *et al.* (2013) Models for open innovation in the pharmaceutical industry. *Drug Discov. Today* 18, 1133–1137
3 Kaitin, K.I. (2010) Deconstructing the drug development process: the new face of innovation. *Clin. Pharmacol. Ther.* 87, 356–361
4 Polishchuk, P.G. *et al.* (2013) Estimation of the size of drug-like chemical space based on GDB-17 data. *J. Comput. Aided Mol. Des.* 27, 675–679
5 Gashaw, I. *et al.* (2011) What makes a good drug target? *Drug Discov. Today* 16, 1037–1043
6 Lazer, D. and Friedman, A. (2007) The network structure of exploration and exploitation. *Admin. Sci. Quart.* 52, 667–694
7 Bentzien, J. *et al.* (2013) We don't know it all: crowd computing to develop and refine predictive *in silico* models. *The Gordon Research Conference on Computer-Aided Drug Design*, West Dover, VT
8 Daemmrich, A. and Bowden, M.E. (2005) Pharma since 1870: a rising drug industry. *Chem. Eng. News* 83, 28–42
9 Howe, J. (2006) The rise of crowdsourcing. Available at: http://archive.wired.com/wired/archive/14.06/crowds.html
10 Sobel, D., ed. (1996) *Longitude: The True Story of a Lone Genius Who Solved the Greatest Scientific Problem of His Time*, Walker and Company, New York
11 Bentzien, J. *et al.* (2013) Crowd computing: using competitive dynamics to develop and refine highly predictive models. *Drug Discov. Today* 18, 472–478
12 Shelley, W. *et al.* (2013) Crowd-sourced soil data for Europe. *Nature* 496, 300
13 Brabham, D.C. (2008) Crowdsourcing as a model for problem solving. *Convergence* 14, 75–90
14 Boudreau, K.J. and Lakhani, K.R. (2013) Using the crowd as an innovation partner. *Harvard Business Rev.* 91, 61–69
15 The ISS Longeron Challenge. Available at: http://www.topcoder.com/iss/longeron/; http://www.nasa.gov/content/coeci/iss-longeron-challenge-completed/#.VBMYNWRdURQ
16 The Asteroid Tracker Challenge Available at: http://www.nasa.gov/content/nasa-ntl-asteroid-tracker-challenge/#.VBMY8mRdURQ; https://community.topcoder.com/longcontest/?module=ViewProblemStatement&rd=16051&pm=13156
17 Anon., . In: http://www.topcoder.com/epa/
18 Anon., . In: http://www.topcoder.com/epa/toxcast/
19 Anon., . In: http://www.topcoder.com/epa/challenge-details/
20 Anon., . In: https://www.gequest.com/c/flight
21 Anon., . In: http://community.topcoder.com/longcontest/?module=ViewProblemStatement&rd=15637&pm=12525
22 Anon., . In: https://www.cog-genomics.org/plink2/credits
23 Lakhani, K. *et al.* (2013) Prize based contests can provide solutions to computational biology problems. *Nat. Biotechnol.* 31, 108–111
24 Markoff, J. (2012) Scientists see advances in deep-learning programs. *New York Times* http://www.nytimes.com/2012/11/24/science/scientists-see-advances-in-deep-learning-a-part-of-artificial-intelligence.html?_r=1&
25 Anon., . In: https://www.crowdanalytix.com/contests/modeling–predict-exacerbation-in-patients-with-respiratory-diseases
26 Novel treatment approaches to cure inflammatory bowel diseases. Available at: https://www.innocentive.com/ar/challenge/9933446
27 Novel implant for the treatment of pelvic organ prolapsed. Available at: https://www.innocentive.com/ar/challenge/9933488
28 A scalable, cost and atom-efficient synthesis of a benzothiazepine. Available at: https://www.innocentive.com/ar/challenge/9933600
29 Shirky, C., ed. (2010) *Cognitive Surplus: Creativity and Generosity in a Connected Age*, The Penguin Press, New York
30 Goetz, T. (2013) How Kaggle is changing how we work. *Atlantic* http://www.theatlantic.com/technology/archive/2013/04/how-kaggle-is-changing-how-we-work/274908/
31 Cooper, S. *et al.* (2010) Predicting protein structures with a multiplayer online game. *Nature* 466, 756–760
32 Khatib, F. *et al.* (2011) Algorithm discovery by protein folding game players. *Proc. Natl. Acad. Sci. U. S. A.* 108, 18949–18953
33 Jorda, J. *et al.* (2014) CrowdPhase: crowdsourcing the phase problem. *Acta Cryst. D* 70, 1538–1548
34 Hong, L. and Page, S.E. (2004) Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proc. Natl. Acad. Sci. U. S. A.* 101, 16385–16389
35 Bohannon, J. (2014) Online video game plugs players into remote-controlled biochemistry lab. *Science* 343, 475
36 Lee, J. *et al.* (2014) RNA design rules from a massive open laboratory. *Proc. Natl. Acad. Sci. U. S. A.* 111, 2122–2127
37 Bell, R.M. and Koren, Y. (2007) Lessons from the Netflix prize challenge, ACM SIGKDD. *Explorations Newsletter* 9, 75–79
38 Tang, J.C. *et al.* (2011) Reflecting on the DARPA Red Balloon Challenge. *Commun. ACM* 54, 78–85
39 Brunsson, N. *et al.* (2012) The dynamics of standardization: three perspectives on standards in organization studies. *Organization Studies* 33, 613–632
40 Gluck, M.E. *et al.* (1987) University–industry relationships in the life sciences: implications for students and post-doctoral fellows. *Res. Policy* 16, 327–336
41 Goldman, M. (2011) Reflections on the innovative medicines initiative. *Nat. Rev. Drug Discov.* 10, 321–322
42 Besnard, J. *et al.* (2014) The Joint European Compound Library: boosting precompetitive research. *Drug Discov. Today* http://dx.doi.org/10.1016/j.drudis.2014.08.014
43 Zdrazil, B. *et al.* (2012) Taking open innovation to the molecular level – strengths and limitations. *Mol. Inf.* 31, 528–535
44 Filippakopoulos, P. *et al.* (2010) Selective inhibition of BET bromodomains. *Nature* 468, 1067–1073
45 Brady, L.S. and Potter, W.Z. (2014) Public–private partnerships to revitalize psychiatric drug discovery. *Expert Opin. Drug Discov.* 9, 1–8
46 Taylor, D. *et al.* (2014) The Pistoia alliance controlled substance compliance service project: from start to finish. *Drug Discov. Today* http://dx.doi.org/10.1016/j.drudis.2014.09.021
47 Goldman, M. (2012) The Innovative Medicines Initiative: a European response to the innovation challenge. *Clin. Pharmacol. Ther.* 91, 418–425
48 Laverty, H. and Goldman, M. (2014) The Innovative Medicines Initiative: a case study for new ways of working in drug R&D. *Future Med. Chem.* 6, 489–491
49 Mullard, A. (2013) European Lead Factory opens for business. *Nat. Rev. Drug Discov.* 12, 173–175
50 Sidders, B. *et al.* (2014) Precompetitive activity to address the biological data needs of drug discovery. *Nat. Rev. Drug Discov.* 13, 83–84
51 Papadaki, M. and Hirsch, G. (2013) Curing consortium fatigue. *Sci. Transl. Med.* 5, 200fs35
52 Hugget, B. (2013) Academic-industry partnerships 2012. *Nat. Biotechnol.* 31, 383
53 Frye, S. *et al.* (2011) US academic drug discovery. *Nat. Rev. Drug Discov.* 10, 409–410
54 Betz, U.A. and Tidona, C.A. (2015) Outcubation—where incubation meets outsourcing. *Nat. Biotechnol.* 33, 20–21
55 Cressey, D. (2011) Traditional drug-discovery model ripe for reform. *Nature* 471, 17–18
56 Robertson, R.M. *et al.* (2014) Open source drug discovery – a limited tutorial. *Parasitology* 141, 148–157
57 Bhardwaj, A. *et al.* (2011) Open source drug discovery – a new paradigm of collaborative research in tuberculosis drug development. *Tuberculosis* 91, 479–486
58 Ioset, J.R. and Chang, S. (2011) Drugs for neglected diseases initiative model of drug development for neglected diseases: current status and future challenges. *Future Med. Chem.* 3, 1361–1371
59 Munos, B. (2010) Can open-source drug R&D repower pharmaceutical innovation? *Clin. Pharmacol. Ther.* 87, 534–536
60 Allarakhia, M. (2014) The successes and challenges of open-source biopharmaceutical innovation. *Expert Opin. Drug Discov.* 9, 459–465
61 Leiter, A. *et al.* (2014) Use of crowdsourcing for cancer clinical trial development. *J. Natl. Cancer Inst.* 106, 1–3
62 Sablinski, T. (2014) Opening up clinical study design to the long tail. *Sci. Trans. Med.* 6, 256ed19
63 Liese, B.H. *et al.* (2014) Development assistance for neglected tropical diseases: progress since 2009. *Int. Health* 6, 162–171
64 Dragojlovic, N. and Lynd, L.D. (2014) Crowdfunding drug development: the state of play in oncology and rare diseases. *Drug. Discov. Today* 19, 1775–1780
65 Kittur, A. *et al.* (2013) The future of crowd work. *CSCW Proceedings of the 2013 Conference on Computer Supported Cooperative Work* pp. 1301–1318
66 Murcko, M.A. and Walters, W.P. (2012) Alpha Shock. *J. Comput. Aided Mol. Des.* 26, 97–102
67 Ingram, P. and Morris, M.W. (2007) Do people mix at mixers? Structure, homophily, and the life of the party. *Admin. Sci. Quart.* 52, 558–585
68 Margolin, A.A. *et al.* (2013) Systematic analysis of challenge-driven improvements in molecular prognostic models for breast cancer. *Sci. Transl. Med.* 5, 181re1

69 Hothorn, T. and Jung, H.H. (2014) RandomForest4Life: a random forest for predicting ALS disease progression. *Amyotroph. Lateral Scler. Frontotemporal Degener.* 15, 444–452
70 Lessl, M. *et al.* (2011) Grants4Targets – an innovative approach to translate ideas from basic research into novel drugs. *Drug Discov. Today* 16, 288–292
71 Dorsch, H. *et al.* (2015) Grants4Targets: an open innovation initiative to foster drug discovery collaborations. *Nat. Rev. Drug Discov.* 14, 74–76
72 Alvim-Gaston, M. *et al.* (2014) Open innovation drug discovery (OIDD): a potential path to novel therapeutic chemical space. *Curr. Top. Med. Chem.* 14, 294–303