Joint NIH-Industry Target Validation Workshop

National Institutes of Health
Building 31, C Wing, Conference Room 10
November 3–4, 2011

Executive Summary

The therapeutics development process is currently experiencing an extremely high attrition rate in late phase clinical trials. Failures due to insufficient efficacy are responsible for 51 percent of Phase II failures[1] and 66 percent of Phase III failures.[2] A major factor implicated in these failures is inadequate target validation, resulting in safe, but ultimately ineffective drugs that fail in clinical trials.

The extraordinary number of potential targets emerging from basic-science discoveries represents an exciting opportunity to address the challenges currently experienced in target validation. For example, genome-wide association studies (GWAS), genome and exome sequencing research, and biological network-based approaches are opening many novel avenues of exploration for therapeutics development. Similarly, efforts to elucidate genetic information from phenotypes associated with or conferring protection against disease are uncovering many potential therapeutic targets. To realize the potential of new scientific discoveries as well as mitigate the current challenges, target validation must be streamlined into a faster, more accurate process to identify the most promising targets and to predict which targets are likely to be biologically relevant and tractable.

While the target validation process has traditionally operated in the private sector, there is growing recognition of the value of collaboration at the precompetitive stage to engage the strengths and expertise of multiple sectors and increase our overall ability to translate new discoveries into clinically useful products. On November 3–4, 2011, a Joint NIH-Industry Workshop on Target Validation was held at the NIH campus to engage experts from industry, government, and academia, to explore the opportunities and challenges in target validation. Specific goals for the meeting included:

- surveying the current state of target validation, including challenges, relevant scientific advances, and other opportunities;
- exploring the advantages and challenges involved in expanding the precompetitive space to enable cross-sector collaboration in target validation efforts; and
- discussing the framework for establishing a Joint NIH-Industry precompetitive target validation consortium.

Major Findings

All workshop participants agreed that a cross-sector consortium, operating in an expanded precompetitive space, would provide the expertise and infrastructure to address current challenges and harmonize the target validation process with scientific advances.

---

[1] Arrowsmith, J. (2011). Trial watch: Phase II failures: 2008–2010. Nat. Rev. Drug Discov, 10(5), 328–329.
[2] Arrowsmith, J. (2011). Trial watch: Phase III and submission failures: 2007–2010. Nat. Rev. Drug Discov., 10(2), 87–87.

Challenges

- *Attrition at all stages of drug development.* Attrition at all stages of drug development is extremely high. Currently, a majority of both Phase II and Phase III trials evaluating novel mechanism of action compounds fail due to inadequate clinical efficacy.

- *Sequencing technology.* Advances in sequencing technology have created a wealth of data on human genetic variation, but also present a challenge in how to interpret findings that will guide the prioritization of leads for promising targets.

- *Early derisking and quick failures*. Patients and advocates, as well as the economy, introduce pressure to choose novel targets (first in class) over improved drugs for existing targets (best in class); however, this comes with a high attrition rate. The participants agreed that new target validation methods are needed to uncover tractable targets more readily and earlier in the therapeutic development process, by taking advantage of data more indicative of therapeutic efficacy. Industry is not averse to failure; however, they do prefer it to occur as early as possible in the drug development pipeline.

- *Avoid repeating work on failed targets*. Participants agreed that an expanded precompetitive space would help avoid duplicative attempts at target validation on failed/difficult targets. Precompetitive collaboration is thought to provide the potential for less duplication of effort and the generation of a greater number of different validated targets to be pursued by all involved.

- *Addressing irreproducibility of published findings*. Industry target validation programs include putative targets identified from published academic findings, in addition to in-house and in-licensed targets. A general rule of thumb often quoted by early venture capital firms is that 50 percent of published results cannot be repeated in the industrial setting. A recently published study of a small collection of validation projects based on academic results indicated that 65 percent could not be replicated in industry validation programs.[3] The participants agreed that a consortium could centralize resources for reproducing results in order to bridge the knowledge/robustness gap between academic findings and validation efforts.

- *Linking sequencing and phenotype to normal biology and disease etiology.* While sequence- and phenotype-driven target qualification represents a more strategic approach to validation, an understanding of the biology underlying the target and disease is critical.

- *Data Management.* Participants agreed on the need for a centralized repository/platform for private sector and public sector data. The need to harmonize a set of core phenotypes was discussed as a way to allow for validation efforts across data sets.

- *Consent.* Participants agreed that a standardized consent form could address problems encountered with reconsent. This has been discussed before, and difficulties experienced in drafting a standard form in previous clinical research should be explored.

- *Experimental design.* Consideration needs to be given to the statistical power when mining data (number of comparisons, significance of effects, and size of cohort).

---

[3] Prinz, F. et al. (2011). Believe it or not: how much can we rely on published data on potential drug targets? Nat. Rev. Drug Discov, 10(9), 712.

Opportunities and Needs

- *Open innovation.* There is a culture change in industry that is moving from close guarding of data to increased sharing across companies and sectors to share risks and rewards in a precompetitive space. Such sharing will also serve to accelerate the understanding of human biology and enhance the ability to aggregate and analyze genetic and phenotypic data.

- *Leveraging existing data.* Existing public and private data sets should be evaluated for new druggable targets in conditions unrelated to the initial research.

    - *Public sector sequencing studies.* NIH has or will support 192 genome and exome sequencing projects representing 68,000 individuals.

    - *Private sector trials.* Industry participants expressed a willingness to share data on consented patients undergoing drug trials (see also *leveraging new data*).

    - One limitation of this approach is the restricted use of data imposed on some studies by the initial informed consents and IRB approvals, thereby preventing application to other disease areas that might benefit but were not included within the scope of the original investigation.

- *Leveraging new data.*

    - *Private sector trials*. Industry participants expressed a willingness to share data on consented research participants enrolled in drug trials.

    - *Non-sequenced populations.* The participants discussed which populations not represented in previous or existing sequencing efforts are important to consider for future investigations.

- *Data repository and analytic platform for data.* Aggregation and analysis of harmonized data sets would provide a powerful resource to identify targets.

- *Areas of opportunity*. Participants discussed advances in sequencing, gene expression, epigenomics, network/omics, perturbagens, multi-cell systems, and tissue engineering, and they discussed how to address specific challenges, opportunities, and needs in these areas.

- *Other stakeholders.* Participants agreed that the roles of other stakeholders, including the Department of Defense, healthcare providers, healthcare payers, and other agencies (FDA, CMS), need to be considered, while moving forward.

- *Training.* Participants discussed the need for more individuals trained in areas aligned with target validation (e.g., physiology, pharmacology, and systems biology).

- *Data access.* The participants agreed that the issue of data access and security will need to be addressed.

- *Precompetitive space.* In defining the precompetitive space, participants agreed that they need to consider the transition to the competitive space and how IP, such as trade secrets and know-how, would be treated or if the space could be completely IP free.

- *Consortium operating parameters.* Participants discussed whether the consortium should be open source or co-exclusive. They agreed that the bounds of the consortium should also be clearly defined.

- *Value proposition.* In considering the establishment of a consortium, participants agreed that generating value propositions for stakeholders in each sector (industry, academia, patients, and government) is integral. Examples include early industry access to data and the ability of consortium participants to have input on the design of new prospective studies that address common goals defined in advance by consortium participants.

- *Pilot projects.* The activation barrier to forming a consortium is high, and developing a few pilots towards proof of concept could generate some quick wins, develop trust, and demonstrate value. These pilots should have defined timelines and milestones that demonstrate progress toward proof of concept within a short time frame (2–3 yrs).

- *Funding.* Funding and in-kind contributions to the consortium should be carefully considered in the context of the framework and governance mechanism, to ensure equitable treatment and shared value for the stakeholders.

Next Steps

- *Identify workstream members from industry, NIH, and academia*. Four workstreams for collaboration were identified during the workshop. Preliminary goals, pilot projects, and important considerations were identified for each. An important first step is for senior pharma participants to assign individuals from their organization to one or more of the workstreams.

    - **Genotype2Phenotype (G2P)**. The goal of this workstream could be to find strong alleles for every gene that might be a good pharma target and evaluate homozygous and heterozygous phenotypes; this effort would entail looking for rare genotypes in unselected populations (e.g., African American ancestry or HMO research networks with electronic health records), as well as existing well-phenotyped cohorts where sequencing efforts are being undertaken (e.g., NIH-funded and other studies). Potential pilots could include longitudinal cohorts not in exome sequencing pipelines; inbred populations and private sector populations not currently in medical pipelines; exploring HMO research networks; and the Million Veterans project.

    - **Phenotype2Genotype (P2G)**. The goal of this workstream could be to find disease/protective alleles in phenotyped individuals. Pilot projects could include existing and future identified cohorts protected from disease (focus on phenotypic outliers), cohorts with disease, non-responders, or other cohorts that may lead to insight into the genetics associated with phenotype.[4]

    - **Information Commons for Biological Function**. The goal of this workstream could be to explore methods for achieving functional integration of diverse biological data, including focused mechanistic studies and multidimensional/high-content data; study and model biological networks as a means to advancing drug discovery and development. This group will require a separate workshop to define needs and opportunities.

[4] The G2P and P2G workstreams have since been consolidated into a single G2P/P2G →F (Function) workstream.

- **Information Commons for Cancer**. A goal of this workstream could be to integrate the currently disparate collections of genotypic, phenotypic, proteomic, immunological, and molecular data on diseases, cell lines, and animal models in a manner allowing for understanding of complex interactions, patterns, and networks. A separate meeting with NCI will be needed to develop this workstream further.

- The workstreams can be encompassed in the following two collaborative spaces:

  - *Learning park*: P2G and G2P involving government, industry, and healthcare. Possible activities include sequencing/genotyping, data aggregation and analysis, development of an informed consent document, evaluating the effect of policies and regulations (e.g., the Health Insurance Portability and Accountability Act, FDA regulatory policies), DNA banking, and phenotyping.

  - *Bio-connectivity Garden:* Information Commons for Biology and Information Commons for Cancer.

- *Determine governance structure for consortium.* This effort will involve a separate working group.