Logical Network Perturbations Guide Precise Target Prioritization for Drug Development

Authors: Saptarshi Sinha[1†], Ella McLaren[1], Madhubanti Mullick[1], Siddharth Singh[2], Brigid S. Boland[2], and Pradipta Ghosh[1, 2†*]

Affiliations:

[1]Department of Cellular and Molecular Medicine, University of California San Diego, 9500 Gilman Drive, La Jolla, CA, 92093, USA.

[2]Department of Medicine, University of California San Diego, 9500 Gilman Drive, La Jolla, CA, 92093, USA.

[†]Institute for Network Medicine, University of California San Diego, 9500 Gilman Drive, La Jolla, CA, 92093, USA.

***Corresponding author.** Email: prghosh@ucsd.edu

Abstract:

Despite advances in artificial intelligence (AI), target-based drug development remains a costly, complex and imprecise process. We introduce F.O.R.W.A.R.D [*Framework for Outcome-based Research and Drug Development*]**, a network-based target prioritization approach and test its utility in the challenging therapeutic area of Inflammatory Bowel Diseases (IBD), which is a chronic condition of multifactorial origin. F.O.R.W.A.R.D leverages real-world outcomes, using a machine-learning classifier trained on transcriptomic data from seven prospective randomized clinical trials involving four drugs. It establishes a molecular signature of remission as the therapeutic goal and computes, by integrating principles of network connectivity, the likelihood that a drug's action on its target(s) will induce the remission-associated genes. Benchmarking F.O.R.W.A.R.D against 210 completed clinical trials on 52 targets showed a perfect predictive accuracy of 100%. The success of F.O.R.W.A.R.D was achieved despite differences in targets, mechanisms, and trial designs. F.O.R.W.A.R.D-driven in-silico phase '0' trials revealed its potential to inform trial design, justify re-trialing failed drugs, and guide early terminations. With its extendable applications to other therapeutic areas and its iterative refinement with emerging trials, F.O.R.W.A.R.D holds the promise to transform drug discovery by generating foresight from hindsight and impacting research and development as well as human-in-the-loop clinical decision-making.**

Introduction

The timeframe for passage of a drug through the three phases of clinical trials on its way to regulatory approval by the Food and Drug Administration (FDA) is approximately 10-12 years (*1*), and costs around $1-1.8 B (*1*). The efficiency of drug development has dramatically decreased since the 1950s; because 90% of the programs fail (*2*), the cost of developing a new drug roughly doubles every nine years, a phenomenon known as "Eroom's Law" (*3, 4*). This is a playful inversion of Moore's Law (a measure of increasing computing efficiency(*5*)), highlighting a paradox that imprecision-related costs of drug discovery escalated despite unprecedented advances in the era of precision medicine (*6, 7*) which witnessed an explosion of artificial intelligence (AI) and computational tools and algorithms designed to sift through 'big data' to extract clinically directive insights. Recently, skeptics have begun questioning if AI has had a measurable impact on drug development (*6, 8, 9*), particularly in three key areas (*10*): (i) *Target discovery*: Many so-called AI-discovered targets were already known through traditional methods and hence, claims of 'new' targets have been questioned. (ii) *Success rates*: AI-driven clinical programs show phase 1 and phase 2 success rates of 39 AI-driven clinical programs from 114 AI-focused companies (~87% and ~40%, respectively (*11*)) similar to those of non-AI programs (~75% and ~50% (*12*)). Phase 3 results are yet to emerge. (iii) *Process innovation*: AI efforts, especially using large language models (LLMs), often replicate traditional methods by combining genetic/transcriptomic data and literature searches to identify disease-target associations, rather than introducing novel approaches. Thus, needs for innovation and precision in target selection are both urgent and unmet, and if fulfilled could address the two main causes of program failure: (i) lack of on-target efficacy (50-60%) or (ii) unmanageable on/off-target toxicity (22-35%) (*13-15*).

Here we sought to develop a framework that uses machine learning approaches to integrate multiple transcriptomic datasets with maximum possible diversity in principled ways. This framework leverages clinical outcomes of past target genes to first define a gene signature of therapeutic remission, and subsequently compute the logical connectivity between remission-associated genes with a drug target gene to inform future target selection and prioritization. Our findings show that the resulting model is both robust and precise and demonstrates the value of its integration as a "module" into the current drug development process. Such integration is expected to enhance precision in target selection and improve accuracy over time, potentially transforming the process of drug discovery.

Results

Study Design and Rationale

To integrate diverse datasets while minimizing the risk of overfitting during machine learning, we adopted a Boolean-logic-based transcriptomic network approach, which relies upon two foundational and intertwined concepts in mathematics and biology (*16*). The Boolean Implication network comprises genes/gene clusters (nodes) interconnected by Boolean Implication Relationships (BIRs; directed edges). BIRs represent all six types of possible gene associations—two symmetric and four asymmetric (*17*); the latter often go unrecognized by conventional methods. BIRs are presumed to reflect fundamental gene regulatory events in the continuum between health and disease (*18-22*). This approach has successfully been used to generate predictive models (a.k.a., disease maps) of sequential changes in gene expression patterns that are fulfilled by all samples (hence, invariant or universal) in the modeled disease (*17, 19, 20, 22-27*). These sequential changes, identified first in a training dataset and prioritized subsequently by machine learning, represent the myriads of continuum states between health and overt disease, manifesting as heterogeneous clinical presentations. The ability to capture invariant gene expression patterns amidst such heterogeneity along a spectrum of severity, disease duration and clinical features allows findings to maintain their relevance across diverse validation datasets. In head-to-head comparisons, gene signatures derived from the Boolean approach outperform other conventional methods (e.g., Differential, Integrated Differential, Bayesian, Hierarchical clustering) for their ability to identify disease-defining gene clusters (*17, 19, 20, 22-27*), indicating a lower risk of overfitting. Perturbations in this network are believed to propagate via fundamental gene regulatory events captured by invariant BIRs, enabling predictable outcomes. While predictable perturbations are expected to overcome the real-world heterogeneity and drive efficacy in phase 3 trials, a focus on the genes helps avoid the biases of functional enrichment analyses (*28, 29*).

Although the Boolean network approach is applicable to any therapeutic area, we chose inflammatory bowel disease (IBD) for three reasons: (i) IBD could benefit from network-modeling approaches because it is a chronic, complex, clinically heterogeneous autoimmune disease of the gut in which diverse components (microbes, genetics, environment and immune response) intersect in elusive ways to culminate in overt disease(*30*). (ii) IBD has a lengthy past and deep current pipeline of drug development programs(*31*) with known and emerging outcomes; they

present an opportunity to test the predictive prowess of any algorithm. Innovation to de-risk phase 2 trials is recognized as an unmet and urgent need in IBD (*32*) because of some very high profile, surprising, and costly phase 3 failures, such as mongersen (anti-sense SMAD7) (*33, 34*), and etrolizumab (β7-integrin) (*35*), among others. (iii) A rigorously validated Boolean-network model of IBD is available(*20*) (**Fig 1**-*Step 1*); developed using ~1500 unique IBD samples, that model identified downregulation of **1284 genes** as a universal molecular phenomenon of IBD which captures the dysfunctional and/or impaired bioenergetics and integrity of the gut barrier (*20*).

We performed machine learning classifier training on the 1284 genes on datasets from 7 cohorts representing real-world clinical trials (n = 322 subjects; **Supplemental Datasheet 1**) to ask and identify what constitutes therapeutic response and if it is universally shared regardless of drug target and class. We then validate findings in 9 cohorts (n = 2336 subjects; **Supplemental Datasheet 1**) representing the diversity of therapeutic modalities, disease activity, presence of inflammation or dysplasia (see **Fig 1**-*Step 2*). A core subset of 34 genes (within the 1284 genes) emerged, whose upregulation is synonymous with response. Next, we formulated the therapeutic goal, i.e., induction of 34 genes through logical perturbations of target genes elsewhere within the network (**Fig 1**-*Step 3*). In creating a framework for drug discovery, we resorted to principled network-based approaches which included a combination of machine learning tools and a suite of statistical learning approaches that estimate the logic, strength, and degree of connectivity of the 34 genes with which the target associates within the network (**Fig 1**-*Step 3*). Finally, we benchmark the framework on 52 targets with known outcomes, compare its predictive accuracy against other approaches and demonstrate its utility in *in-silico* phase 0 trials on custom cohorts (**Fig 1**-*Step 4*).

Identification of a molecular signature of clinical remission and therapeutic goal

We used as starting point the 1284 genes whose consistent downregulation in all IBD samples was found to be indicative of barrier defect (*20*) (**Fig 2A**-*left*). Because a compromised epithelial barrier is the most widely recognized indicator/predictor of disease relapse, response and remission (*36-41*), we asked if certain aspects of the compromised barrier determine remission. We trained the 1284-gene signature by assessing each gene for its ability to classify pre-treatment samples from responders and non-responders and do so consistently across 7 prospective randomized placebo-controlled clinical trials of 4 FDA-approved biologics (**Fig 2A**-*middle*).

Application of a stringent cutoff during such outcome-based training (mean AUC 0.70 and standard deviation of 0.07) yielded a refined core subset of 34 genes (**Fig 2A**-*right*). These 34 genes classified responders and non-responders reliably across all training cohorts (**Fig 2B**), regardless of the biologic used or their mechanism(s) of action. Multivariate analyses on all training cohorts in which >/= 3 variables were annotated showed that the 34 genes positively correlated with response (**Fig 2C-E**), as determined by endoscopic and histologic documentation of mucosal healing, an objective and stringent desirable endpoint shared by most current IBD clinical trials.

Validation studies showed that the 34-gene signature classified responders from non-responders regardless of the type (biologics or cell-based) or mechanism of action (anti-inflammatory, anti-trafficking, or regenerative) of the therapeutic intervention (**Fig 2F**-*top*). The 34 gene signature of barrier dysfunction could also classify the samples across other validation cohorts (**Fig 2F**-*bottom*), and its downregulation was a shared feature among those with disease, especially active disease, the presence of inflammation, and those at risk of neoplasia. We explored the relevance of the 34 gene signature in the heterogeneous complex cohorts in which samples were annotated with variables related to clinical outcomes. Multivariate analyses in those cohorts confirmed that the 34 genes are consistently and inversely correlated with the presence of histological inflammation (**Fig 2G**) and endoscopic severity (**Fig 2H-I**).

A composite score analysis for the levels of expression of the 34 gene signature revealed that its upregulation during therapy is a feature of therapeutic response (**Fig 2J**). Similar comparative analyses of composite scores revealed that the 34 genes were suppressed regardless of disease subtype, i.e., both Ulcerative colitis (UC) and Crohn's disease (CD) (2 different cohorts; **Fig 2K-L**). Comparisons of the composite scores before and after therapy showed that its upregulation was significantly higher among responders (**Fig 2M**) compared to non-responders (**Fig 2N**) and that this pattern generally held true regardless of disease subtype (**Fig 2O-P**). Gene ontology analyses using a variety of approaches showed that core processes concerning cellular bioenergetics such as mitochondrial electron transport chain, cytochrome C oxidase activity and sulfur metabolism and detoxification of hydrogen sulfide (H2S) are significantly enriched in the 34 genes (**Fig S1A-D**).

These findings suggest that the 34-genes are a molecular signature of the gut's bioenergetic state, which is a key indicator of clinical response: Responders have higher baseline

expression and continue to induce the 34 genes during treatment, whereas non-responders have lower baseline expression and/or fail to induce it during treatment. Most importantly, the 34 gene signature tracked the endoscopic and histologic proof of mucosal healing at the end of treatment, which is an increasingly used endpoint in clinical trials. Thus, pharmacologic induction of the 34 genes emerged as a therapeutic goal.

A Framework of Network-based Rules for Target Prioritization

We asked if the 34 genes were associated (i.e., connected via BIRs) with target genes of drugs used in the training cohorts such that their induction could be enabled by target perturbation. We created a comprehensive list of genes encoding the protein targets of those drugs (**Supplemental Datasheet 2**). For each target gene, we systematically computed two statistical values: (i) therapeutic index (TI); and (ii) network connectivity index (CI) and an overall likelihood of approval index (LoAI) for each drug (**Fig 3A**). To compute these, we used the same dataset that was used to build the original Boolean implication network model(*20*). This dataset provides key advantages such as being of high-quality, as determined by a >90% of reads that align to the reference genome or transcriptome and represents the largest cohort comprised of full thickness (surgically resected) samples from patients that are naïve to any biologic agent (*42*).

Therapeutic index reflects both the logic/type and the statistical strength of the BIR of any query gene (i.e., target gene, T) with any of the 34 genes (i.e., reference gene, R) within an IBD transcriptome dataset (see detailed *Methods* and **Fig S2A-E**). TI informs if a target's transcripts should be upregulated (with agonists) or downregulated (with antagonists) to get the desired outcome, i.e., induction of as many of the 34 reference genes (**Fig 3B**). The strength of association informs how likely it is that the manipulation of the target gene will induce (desired/intended effect) or suppress (unintended effect) the expression of the reference genes of molecular remission.

Connectivity index refers to the fraction of 34 reference genes to which any target gene is connected fulfilling a stringent cutoff of < 0.075 TI. By a simple rule of majority, targets with sparse connectivity (<50% of 34 genes) are expected to perform poorly, and those with dense connectivity (>/= 50% of 34 genes) (**Fig 3C**). Simple majority protocol was favored because of its proven ability to reach a consensus and control network behavior within fully connected large networks, especially in those where multiple factors communicate in unreliable ways (*43*).

As many drugs target more than one gene, we derived a drug-specific LoAI; it is the cumulative non-overlapping union of the CI of each drug's target genes such that each of the 34

reference genes is included only once, resulting in a comprehensive score without redundancy (see example; **Fig 3D**).

For each of the 4 drugs in the training cohorts that crossed the high bar of phase 3 trials and subsequent regulatory approval (**Fig 3E**-*yellow dots*), key genes they antagonistically targeted provided >50% connectivity with the 34 reference genes (**Fig 3E**-top, *black*). By contrast, the genes that are antagonistically targeted by the failed drug programs, e.g., *SMAD7*, *MAdCAM-1*, *ITGB7* and *IL36*, provided sparse connectivity with the reference genes (**Fig 3E**-bottom, *grey*). It is noteworthy that LoAI accounts for multi-target drugs with inconsistent connectivity, wherein intended vs unintended effects on 34 genes of remission completely negate each other (**Fig S3A-B**, showcasing the terminated trial of pan-HDAC inhibitor givinostat; NCT00792740). The CI of each drug/target pair is presented in **Supplemental Datasheet 2**. Based on this connectivity pattern, the targets of several ongoing clinical trials (**Fig 3E**-red dots) clustered with successful drug programs, whereas others clustered with known failed programs.

Because agonists are yet to succeed in phase 3 trials, we assessed all failed targets of the past alongside candidates in ongoing trials (**Fig 3F**). We also included in the analysis two agonists that have emerged from the Boolean transcriptomics approach, the PRKAB1-agonist PF-06409577(*20*) and the balanced PPARα/γ-dual agonist PAR5359 (*26*). None of the agonists of the past crossed the bar of 50% connectivity.

These findings indicate that drugs successfully advancing through phase 3 trials exhibit greater connectivity—measured in terms of strength, degree, and logical relationships—with the core 34 reference genes within the IBD network. Because this framework estimates the connectivity of potential target genes with the 34 reference genes that were derived from real-world outcomes, we hereafter refer to it as *F.O.R.W.A.R.D*: Framework for Outcome-based Research and Drug Development.

Benchmarking F.O.R.W.A.R.D.

Next we asked how F.O.R.W.A.R.D's predictive accuracy compares against other conventional approaches that are currently used for target identification (**Fig 4A-C**). Differential expression analysis (DEA) is the most widely used approach and the current gold standard. Generative-AI models for target discovery, e.g., PandaOmics (*28*) (used by Insilico Medicine) also use DEA as its core approach. A Bayesian method, which offers distinct advantages (*44*), has been used to integrate the genomic, transcriptomic and clinical data from patients with IBD and create functional genomics predictive network models for IBD (*42*). This effort identified a set of 133 key driver genes (KDGs) and validated a subset of the top-ranked 12 KDGs that regulate processes

central to IBD (*42*). We also included a 74-gene signature that was identified using artificial neural networks (ANN) as potential drivers of IBD (*45*). Limited overlaps were noted among the genes of the 4 signatures (**Fig 4A**). To determine the distribution and basic properties of each gene signature and its relationship with response (the variable of interest in this case), we conducted univariate analyses on the clinical trial datasets (**Fig 4D**). F.O.R.W.A.R.D demonstrated consistency across them all, showing significant positive correlations with response in 8/9 cohorts. The 133-gene Bayesian KDGs reached significance in 7/9 cohorts but showed inconsistency in correlation with response (negatively correlated in all cases, except in GSE76331; **Fig 4D-***3rd plot from the left*). The top-ranked 12-gene KDG subset, DEA and ANN underperformed. The superset of 1284 Boolean network genes from which F.O.R.W.A.R.D was derived, was superior to other conventional approaches (reaching significance in 7/9 cohorts) but underperformed when compared to its outcome-trained 34-gene subset.

A list of successful targets was curated from FDA.gov and augmented further through literature search for international approvals [e.g., Filgotinib in Japan, Great Britain and the EU, AJM300 (Carogra®) (*46*) in Japan]. A list of failed targets from phase 2-3 was curated from *ClinicalTrials.gov*, and the European Medicines Agency (EMA) Clinical Trials Register and augmented further using Industry Press releases. To focus on efficacy, early-development programs that were terminated for concerns of safety were excluded, which led to the identification of 52 targets tested in 210 clinical trials on adult patients with IBD (**Fig 4B-C**). F.O.R.W.A.R.D outperformed all other approaches in picking successes (LoAI>/= 0.5) from failures (LoAI <0.5) (Fisher's exact test: 6.3e-11; **Fig 4E**). Notably, mongersen (GED-301), that reached a primary end point in a double-blind, placebo-controlled, phase 2 trial (*46*) and yet, failed phase 3, scored 0 (**Fig 4E**). Similarly, the anti-β7 integrin antibody etrolizumab, which showed promising phase 2 data (*47*), but failed to get regulatory approval after five phase 3 trials (*35, 48-52*), scored 0.029 (**Fig 4E**). Cobitolimod, the TLR9 agonist, which cleared a randomized, double-blind, placebo-controlled phase 2 trial(*53*) but was abandoned during phase 3 citing lack of efficacy (NCT04985968), scored 0 (**Fig 4D**). Ontamalimab, which targets MadCAM-1, and underwent lengthy evaluation in six phase 3, multicenter, randomized, double-blind, placebo-controlled clinical trials (*54*) before its termination in 2020, scored 0 (**Fig 4E**). Thus, the FORWARD model, which relies on the 34 gene signature of therapeutic response, appears to capture the potential for regulatory approval regardless of therapeutic modality.

F.O.R.W.A.R.D-driven in silico trials

F.O.R.W.A.R.D can be integrated in early clinical phases to vet and prioritize targets as single or multiple genes (**Fig 5A**-*left*), or for expanded screens of an entire target class (such as kinome,

GPCRome, etc.). Because F.O.R.W.A.R.D can 'learn' from emerging studies with known outcomes, it can be integrated in the late clinical phases (outcomes-based learning; **Fig 5A**-*right*). Besides these obvious points of integration, we tested the ability of F.O.R.W.A.R.D to inform clinical trial design and cohort selection through *in silico* trials. Although the original IBD network was built and the 1284 gene signature of barrier integrity was validated using ~2700 unique but diverse human IBD samples (**Fig 1**-*step 1*), network connectivity of target gene to the 34 gene molecular signature of response was calculated only in a cohort that is naïve to any biologics (**Fig 3E-F**). We asked how network connectivity and hence, predicted efficacy of drugs may be influenced by prior exposure to biologics.

We computed the LoAI of all drugs in 3 *in silico* trials: (i) a cohort of biologic naïve, which remains the original cohort used for building the Boolean Implication network (*20*) (**Fig 5B**); (ii) a large complex cohort that resembles real-world scenario, in which ~75% subjects have late-stage disease and have been exposed to biologics(*55*) (**Fig 5C**); and finally, (iii) a cohort (*56*) in which all subjects have been acutely exposed to two different classes of biologics, infliximab (4-6 weeks after a single dose) and vedolizumab (6, 12 and 52 weeks post-treatment; from two phase 3 trials, GEMINI I and LTS). Comparative analyses of the LoAI of each drug across these trials revealed striking changes in the predicted efficacy of drugs in bio-exposed complex cohorts compared to the bio-naïve setting (**Fig 5B** *vs* **Fig 5C-D**). This is consistent with the widely accepted belief that dynamic rewiring of the colon network ensues during disease progression (*57*), which is believed to be one of the drivers of emergence of therapeutic resistance during the long-term use of any therapeutic agent. Antagonists of *JAK, TYK2, PIKfyve*, and LIGHT are expected to gain efficacy in the complex chronic setting, whereas vedolizumab (α4β7), ozanimod (S1PR1) and etrasimod (S1PR1/4/5) are predicted to drop in efficacy. Ustekinumab, risankizumab and mirikizumab (IL12/23) were predicted to maintain efficacy across the board. Existing evidence suggests that at least some of these predictions are true (see *Discussion*). Perhaps the most unexpected finding is that many failed drugs are predicted to succeed if they were to be re-trialed in the setting of acute exposure to either infliximab or vedolizumab (**Fig 5D**), notable examples of which include: antagonists of CTLA4, IL17A, MMP9 and IL10. In the absence of clinical trials, these beneficial predictions of sequential therapy can neither be supported, nor challenged yet with evidence. The pan-HDAC inhibitor, givinostat, continued to perform poorly in all three *in silico* trials, and in fact, it is predicted to have unintended effects (e.g., possible on-treatment flares) in the most complex cohort (**Fig S3B-D**).

Furthermore, we asked if F.O.R.W.A.R.D could indicate superiority in head-to-head comparisons. Between the three IL12/23 inhibitors, risankizumab and mirikizumab has been found to outperform ustekinumab in real-world trials, both in indirect(*58, 59*) and head-to-head

direct comparisons (phase 3 SEQUENCE (*60*) trial). We noted that although all of them pass the LoAI threshold of efficacy, risankizumab and mirikizumab have a slightly higher LoAI than ustekinumab in the bio-naïve cohort (**Fig 5B**), and higher overlaps in reference gene connectivity in the bio-exposed cohort, which can be measured by calculating cumulative connectivity (C_T) (**Fig 5D-E**).

These findings demonstrate how F.O.R.W.A.R.D can be seamlessly integrated with the current workflow in any therapeutic area, and how its impact on target identification/prioritization and target-specific trial design (including re-trials for failed programs or rationalize head-to-head comparisons) can be realized, tested and improved.

Discussion:

The major innovation we present here is a computational framework designed to enhance precision in target discovery. The key outcomes of this study highlighting the four distinguishing features of the framework and their potential impact are summarized below.

First, F.O.R.W.A.R.D enables target prioritization based on an objectively defined therapeutic goal at a molecular level, which is, to upregulate a 34-gene signature of remission by perturbing strongly associated neighboring genes with targeted drugs. In doing so, F.O.R.W.A.R.D preserves the importance of the target gene(s) and raises the odds of identifying new targets while reducing biases and avoiding the introduction of biases in functional enrichment analyses (*28, 29*). Validation in diverse IBD datasets confirmed the 34-gene signature's predictive power, demonstrating its potential as a robust biomarker for the intactness of the gut barrier, therapeutic response and disease remission. Gene ontology analyses linked the 34 genes to cellular bioenergetic processes, covering almost every core mitochondrial function. This is not surprising given that perturbed mitochondrial genes (*61*) and functions (biogenesis, quality control and metabolism) are a feature of intestinal inflammation (*62*), promotes dysbiosis (*63*), and whose reversal has therapeutic potential in IBD (*64*). A comprehensive literature search failed to identify GWAS evidence for these 34 genes either as contributors of IBD risk or as associations of IBD features/severity, suggesting that epigenetics may be in play. Examples of such epigenetic control of bioenergetics that impact IBD progression and remission exist (*65-70*).

Second, benchmarking against conventional approaches like DEA, ANN and Bayesian methods demonstrate F.O.R.W.A.R.D.'s superiority in identifying successful drug programs. A notable surprise was the degree of precision and objectivity with which it pinpointed some of the most unexpected failures of the past. If applied, objective scores from F.O.R.W.A.R.D would warrant an early termination for all of them and more (see **Fig 4E**). As for unexpected results on

ongoing trials, the most notable was TL1A, which has completed three phase 2 studies (DIONE, TUSCANY(*71*) and TUSCANY II (*72*)). Among these, the DIONE and the TUSCANY II studies incorporated a companion biomarker to enhance treatment outcome reports in biomarker-positive patients. F.O.R.W.A.R.D. predicts TL1A will succeed if phase 3 studies are evaluated for efficacy selectively among those positive for biomarker. We conclude that by recognizing programs that are worthy of early termination or abandonment, F.O.R.W.A.R.D. can cut the costs associated with later stages of drug development. By de-risking target selection, F.O.R.W.A.R.D may reduce the trade-off between choosing high-confidence and novel targets.

Third, we show that F.O.R.W.A.R.D. can be integrated at critical decision points in the drug development pipeline. With rising stringency of trial designs (*73*), recruiting enough patients (biologic naïve) is identified as the biggest challenge(*74*), and hence, innovation in study design is identified as the greatest opportunity (*75*). F.O.R.W.A.R.D's predictive power can be implemented in phase '0' *in silico* trials on custom cohorts (severity, disease subtype, disease stage, exposure to biologics, etc.) to inform trial design. Although it is difficult to assess the usefulness of each of these predictions in the absence of trials that have tested the same, some of the findings are consistent with real world outcomes. For example, Tofacitinib has been associated with superior effectiveness outcomes compared with vedolizumab in anti-TNF-exposed patients (*76*) and anti-IL12/23 therapeutics maintain similar efficacy across the board (*77, 78*). A loss of efficacy with vedolizumab in patients exposed to anti-TNF or with complex severe disease has been observed (*79, 80*). We conclude that by recognizing which failed drugs might succeed if re-trialed in specific cohorts or in specific sequences, costly ventures of the past could emerge as valuable assets. Using the IL12/23 antagonists as examples, we show that drugs that pass F.O.R.W.A.R.D's LoAI criteria, could be additionally assessed by cumulative connectivity (C_T) to gauge performance in head-to-head trials (see **Fig 5E**).

Fourth, F.O.R.W.A.R.D's predictive powers can be enhanced through iterative refinement over time through further training on datasets that emerge from late clinical phases. Such training on both failed and successful programs is expected to iteratively refine the 34-gene molecular signature of response, which can be retrained to adjust cut-offs in TI without altering CI that requires a simple majority. Although one such 'learning' model exists (*81*), it relies for target identification on conventional DEA and functional enrichment analyses (*28, 29*); the latter is prone to biases (*29, 82-88*), i.e., skewed towards richly annotated genes instead of those with the strongest molecular data (*82*). The lead drug program from this model prioritized an anti-fibrotic

target, TINK (TRAF2- and NCK-interacting kinase) and engaged it with an AI-optimized small molecule (INS018_055) for combating Idiopathic Pulmonary Fibrosis (IPF) (*89*). The program is in the early stages of phase 2, and outcomes have yet to emerge.

Limitations

F.O.R.W.A.R.D can address target novelty focused exclusively on high efficacy; but cannot address druggability (potency) and toxicity, two other crucial factors in target selection. This is why it fails to predict why the IL12(p40)/23 antagonist ustekinumab passed and yet, briakinumab failed due to severe adverse effects (*90*). F.O.R.W.A.R.D can, however, compute off-target unintended effects, by using a similar approach as demonstrated for pan-HDAC inhibitors. F.O.R.W.A.R.D cannot compute the boost in efficacy that results from the use of companion biomarker(s) [e.g., in TLA1-centric trials], unless it is a transcriptome-based biomarker that can be tested through *in silico* trials. F.O.R.W.A.R.D also cannot compute the impact of strategic decisions, e.g., the IL12/23(p19) antagonists risankinumab and mirikizumab pass, but brazikumab was abandoned citing trial delays. Because F.O.R.W.A.R.D's predictive abilities rely upon BIRs that capture the underlying fundamental gene regulatory events (*16, 17*), experimental confirmation that intended mechanisms of action as agonist vs antagonist leads to up- and downregulation, respectively, of the target transcript is warranted. This is particularly important when studying high-confidence targets for which receptor activity and the abundance of transcripts dissociates in non-intuitive ways (*91*), with unpredictable transcript patterns in low vs high doses (*92*), e.g., the terminated IBD programs, Fingolimod (NCT01375179), a S1PR1/5 inhibitor or Ozanimod (the YELLOWSTONE program (*93*)). In such cases, F.O.R.W.A.R.D warrants retrials at different doses that achieve the molecular therapeutic goal.

Conclusions:

The F.O.R.W.A.R.D. framework introduces a novel, data-driven approach to target selection in drug development, promising enhanced precision and predictive accuracy. By integrating real-world clinical outcomes with principles of network biology, it marks a significant detour from the well-trodden paths. By retaining human interpretability F.O.R.W.A.R.D. ensures that model outputs are easily understood, interpreted, and actionable; this is in sharp contrast to the ‘black box’ opacity of deep learning models that exclude humans in the loop. Moreover, the framework is simple enough to be reproducibly implemented and seamlessly integrated as a module into the current drug discovery process. These features are expected to enable immediate and widespread adoption, augmenting conventional approaches and enhancing precision in target

selection. F.O.R.W.A.R.D. demonstrates that hindsight can indeed be leveraged to create foresight, as predictable failures can be avoided. Hence, it presents a unique framework for leveraging outcome knowledge to reduce uncertainty.

References and Notes

1. S. M. Paul *et al.*, How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov* **9**, 203-214 (2010).

2. D. Sun, W. Gao, H. Hu, S. Zhou, Why 90% of clinical drug development fails and how to improve it? *Acta Pharm Sin B* **12**, 3049-3062 (2022).

3. J. W. Scannell, A. Blanckley, H. Boldon, B. Warrington, Diagnosing the decline in pharmaceutical R&D efficiency. *Nat Rev Drug Discov* **11**, 191-200 (2012).

4. L. Derek, Science, Ed. (2012).

5. R. R. Schaller, Moore's law: past, present and future. *IEEE spectrum* **34**, 52-59 (1997).

6. L. Yu, Artificial Intelligence in Molecular Medicine. *N Engl J Med* **389**, 1251-1252 (2023).

7. F. W. Pun, I. V. Ozerov, A. Zhavoronkov, AI-powered therapeutic target discovery. *Trends in Pharmacological Sciences*, (2023).

8. L. Dere. (AAAS, Science, 2024).

9. L. Dere. (AAAS, Science, 2023).

10. L. Dere. (Science, 2024).

11. M. Kp Jayatunga, M. Ayers, L. Bruens, D. Jayanth, C. Meier, How successful are AI-discovered drugs in clinical trials? A first analysis and emerging lessons. *Drug Discov Today* **29**, 104009 (2024).

12. T. Takebe, R. Imai, S. Ono, The Current Status of Drug Discovery and Development as Originated in United States Academia: The Influence of Industrial and Academic Collaboration on Drug Discovery and Development. *Clin Transl Sci* **11**, 597-606 (2018).

13. J. Arrowsmith, Phase II failures: 2008–2010. *Nature Reviews Drug Discovery* **10**, 328-329 (2011).

14. J. Arrowsmith, P. Miller, Phase II and Phase III attrition rates 2011–2012. *Nature Reviews Drug Discovery* **12**, 569-569 (2013).

15. R. K. Harrison, Phase II and phase III failures: 2013–2015. *Nature Reviews Drug Discovery* **15**, 817-818 (2016).

16. D. Sahoo, The power of boolean implication networks. *Front Physiol* **3**, 276 (2012).

17. D. Sahoo, D. L. Dill, A. J. Gentles, R. Tibshirani, S. K. Plevritis, Boolean implication networks derived from large scale, whole genome microarray datasets. *Genome biology* **9**, R157 (2008).

18. D. Sahoo *et al.*, MiDReG: a method of mining developmentally regulated genes using Boolean implications. *Proc Natl Acad Sci U S A* **107**, 5732-5737 (2010).

19. P. Ghosh *et al.*, Machine learning identifies signatures of macrophage reactivity and tolerance that predict disease outcomes. *EBioMedicine* **94**, 104719 (2023).

20. D. Sahoo *et al.*, Artificial intelligence guided discovery of a barrier-protective therapy in inflammatory bowel disease. *Nat Commun* **12**, 4246 (2021).

21. D. Vo, P. Ghosh, D. Sahoo, Artificial intelligence-guided discovery of gastric cancer continuum. *Gastric Cancer* **26**, 286-297 (2023).

22. P. Ghosh *et al.*, AI-assisted discovery of an ethnicity-influenced driver of cell transformation in esophageal and gastroesophageal junction adenocarcinomas. *JCI Insight* **7**, (2022).

23. P. Dalerba *et al.*, Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat Biotechnol* **29**, 1120-1127 (2011).

24. P. Dalerba *et al.*, CDX2 as a Prognostic Biomarker in Stage II and Stage III Colon Cancer. *N Engl J Med* **374**, 211-222 (2016).

25. S. Sinha *et al.*, Network-guided Therapeutics to Differentiate Colorectal Cancer Stem Cells. *bioRxiv*, (2023).

26. G. D. Katkar *et al.*, Artificial intelligence-rationalized balanced PPARα/γ dual agonism resets dysregulated macrophage processes in inflammatory bowel disease. *Commun Biol* **5**, 231 (2022).

27. P. E. Zage *et al.*, Identification of a novel gene signature for neuroblastoma differentiation using a Boolean implication network. *Genes Chromosomes Cancer* **62**, 313-331 (2023).

28. P. Kamya *et al.*, PandaOmics: An AI-Driven Platform for Therapeutic Target and Biomarker Discovery. *J Chem Inf Model* **64**, 3961-3969 (2024).

29. F. W. Pun *et al.*, Identification of Therapeutic Targets for Amyotrophic Lateral Sclerosis Using PandaOmics - An AI-Enabled Biological Target Discovery Platform. *Front Aging Neurosci* **14**, 914017 (2022).

30. C. Abraham, J. H. Cho, Inflammatory bowel disease. *N Engl J Med* **361**, 2066-2078 (2009).

31. D. C. Baumgart, C. Le Berre, Newer Biologic and Small-Molecule Therapies for Inflammatory Bowel Disease. *N Engl J Med* **385**, 1302-1315 (2021).

32. N. M. Noor, T. Raine, Innovations to improve the efficiency of phase II IBD clinical trials. *Nat Rev Gastroenterol Hepatol* **20**, 555-556 (2023).

33. B. E. Sands *et al.*, Mongersen (GED-0301) for Active Crohn's Disease: Results of a Phase 3 Study. *Am J Gastroenterol* **115**, 738-745 (2020).

34. G. Monteleone, C. Stolfi, I. Marafini, R. Atreya, M. F. Neurath, Smad7 Antisense Oligonucleotide-Based Therapy in Crohn's Disease: Is it Time to Re-Evaluate? *Mol Diagn Ther* **26**, 477-481 (2022).

35. W. J. Sandborn *et al.*, Etrolizumab as induction and maintenance therapy in patients with moderately to severely active Crohn's disease (BERGAMOT): a randomised, placebo-controlled, double-blind, phase 3 trial. *Lancet Gastroenterol Hepatol* **8**, 43-55 (2023).

36. R. D'Incà *et al.*, Intestinal permeability test as a predictor of clinical course in Crohn's disease. *Am J Gastroenterol* **94**, 2956-2960 (1999).

37. R. Kiesslich *et al.*, Local barrier dysfunction identified by confocal laser endomicroscopy predicts relapse in inflammatory bowel disease. *Gut* **61**, 1146-1153 (2012).

38. W. Fries, A. Belvedere, S. Vetrano, Sealing the broken barrier in IBD: intestinal permeability, epithelial cells and junctions. *Curr Drug Targets* **14**, 1460-1470 (2013).

39. J. Florholmen, Mucosal healing in the era of biologic agents in treatment of inflammatory bowel disease. *Scand J Gastroenterol* **50**, 43-52 (2015).

40. J. Chang *et al.*, Impaired Intestinal Permeability Contributes to Ongoing Bowel Symptoms in Patients With Inflammatory Bowel Disease and Mucosal Healing. *Gastroenterology* **153**, 723-731.e721 (2017).

41. L. Shen, L. Su, J. R. Turner, Mechanisms and functional implications of intestinal barrier defects. *Dig Dis* **27**, 443-449 (2009).

42. L. A. Peters *et al.*, A functional genomics predictive network model identifies regulators of inflammatory bowel disease. *Nat Genet* **49**, 1437-1449 (2017).

43. R. Tamir, A. Livshits, Y. Shadmi, Simple Majority Consensus in Networks with Unreliable Communication. *Entropy (Basel)* **24**, (2022).

44. M. Stephens, D. J. Balding, Bayesian statistical methods for genetic association studies. *Nat Rev Genet* **10**, 681-690 (2009).

45. N. H. Nguyen *et al.*, Machine Learning-based Prediction Models for Diagnosis and Prognosis in Inflammatory Bowel Diseases: A Systematic Review. *J Crohns Colitis* **16**, 398-413 (2022).

46. S. Dhillon, Carotegrast Methyl: First Approval. *Drugs* **82**, 1011-1016 (2022).

47. S. Vermeire *et al.*, Etrolizumab as induction therapy for ulcerative colitis: a randomised, controlled, phase 2 trial. *Lancet* **384**, 309-318 (2014).

48. L. Peyrin-Biroulet *et al.*, Etrolizumab as induction and maintenance therapy for ulcerative colitis in patients previously treated with tumour necrosis factor inhibitors (HICKORY): a phase 3, randomised, controlled trial. *Lancet Gastroenterol Hepatol* **7**, 128-140 (2022).

49. S. Danese *et al.*, Etrolizumab versus infliximab for the treatment of moderately to severely active ulcerative colitis (GARDENIA): a randomised, double-blind, double-dummy, phase 3 study. *Lancet Gastroenterol Hepatol* **7**, 118-127 (2022).

50. D. T. Rubin *et al.*, Etrolizumab versus adalimumab or placebo as induction therapy for moderately to severely active ulcerative colitis (HIBISCUS): two phase 3 randomised, controlled trials. *Lancet Gastroenterol Hepatol* **7**, 17-27 (2022).

51. S. Vermeire *et al.*, Etrolizumab for maintenance therapy in patients with moderately to severely active ulcerative colitis (LAUREL): a randomised, placebo-controlled, double-blind, phase 3 study. *Lancet Gastroenterol Hepatol* **7**, 28-37 (2022).

52. W. J. Sandborn *et al.*, Etrolizumab for the Treatment of Ulcerative Colitis and Crohn's Disease: An Overview of the Phase 3 Clinical Program. *Adv Ther* **37**, 3417-3431 (2020).

53. R. Atreya *et al.*, Cobitolimod for moderate-to-severe, left-sided ulcerative colitis (CONDUCT): a phase 2b randomised, double-blind, placebo-controlled, dose-ranging induction trial. *Lancet Gastroenterol Hepatol* **5**, 1063-1075 (2020).

54. S. Vermeire *et al.*, Efficacy and Safety of the Anti-mucosal Addressin Cell Adhesion Molecule-1 Antibody Ontamalimab in Patients with Moderate-to-Severe Ulcerative Colitis or Crohn's Disease. *J Crohns Colitis* **18**, 708-719 (2024).

55. C. Argmann *et al.*, Biopsy and blood-based molecular biomarker of inflammation in IBD. *Gut* **72**, 1271-1287 (2023).

56. I. Arijs *et al.*, Effect of vedolizumab (anti-α4β7-integrin) therapy on histological healing and mucosal gene expression in patients with UC. *Gut* **67**, 43-52 (2018).

57. C. S. Smillie *et al.*, Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. *Cell* **178**, 714-730.e722 (2019).

58. M. Dubinsky *et al.*, Matching-Adjusted Indirect Comparison Between Risankizumab and Ustekinumab for Induction and Maintenance Treatment of Moderately to Severely Active Crohn's Disease. *Adv Ther* **40**, 3896-3911 (2023).

59. V. Jairath *et al.*, OP35 Efficacy of mirikizumab in comparison to ustekinumab in patients with moderate to severe Crohn’s disease: Results from the phase 3 VIVID 1 study. *Journal of Crohn's and Colitis* **18**, i62-i64 (2024).

60. L. Peyrin-Biroulet *et al.*, S1 Risankizumab Versus Ustekinumab in Patients With Moderate to Severe Crohn’s Disease: Results From the Phase 3B SEQUENCE Trial. *Official journal of the American College of Gastroenterology| ACG* **118**, S1 (2023).

61. J. Chen *et al.*, Multi-omic insight into the molecular networks of mitochondrial dysfunction in the pathogenesis of inflammatory bowel disease. *EBioMedicine* **99**, 104934 (2024).

62. G. T. Ho, A. L. Theiss, Mitochondria and Inflammatory Bowel Diseases: Toward a Stratified Therapeutic Intervention. *Annu Rev Physiol* **84**, 435-459 (2022).

63. A. Peña-Cearra *et al.*, Mitochondrial dysfunction promotes microbial composition that negatively impacts on ulcerative colitis development and progression. *NPJ Biofilms Microbiomes* **9**, 74 (2023).

64. N. L. Mancini *et al.*, Perturbed Mitochondrial Dynamics Is a Novel Feature of Colitis That Can Be Targeted to Lessen Disease. *Cell Mol Gastroenterol Hepatol* **10**, 287-307 (2020).

65. N. T. Ventham *et al.*, Integrative epigenome-wide analysis demonstrates that DNA methylation may mediate genetic risk in inflammatory bowel disease. *Nat Commun* **7**, 13507 (2016).

66. J. Cosin-Roger *et al.*, Hypoxia ameliorates intestinal inflammation through NLRP3/mTOR downregulation and autophagy activation. *Nat Commun* **8**, 98 (2017).

67. P. S. Karatzas, G. J. Mantzaris, M. Safioleas, M. Gazouli, DNA methylation profile of genes involved in inflammation and autoimmunity in inflammatory bowel disease. *Medicine (Baltimore)* **93**, e309 (2014).

68. M. A. Halili, M. R. Andrews, M. J. Sweet, D. P. Fairlie, Histone deacetylase inhibitors in inflammatory disease. *Curr Top Med Chem* **9**, 309-319 (2009).

69. C. Shi *et al.*, MicroRNA-21 knockout improve the survival rate in DSS induced fatal colitis through protecting against inflammation and tissue injury. *PLoS One* **8**, e66814 (2013).

70. J. Hu *et al.*, miR-155: An Important Role in Inflammation Response. *J Immunol Res* **2022**, 7437281 (2022).

71. S. Danese *et al.*, Anti-TL1A Antibody PF-06480605 Safety and Efficacy for Ulcerative Colitis: A Phase 2a Single-Arm Study. *Clin Gastroenterol Hepatol* **19**, 2324-2332.e2326 (2021).

72. L. Stephani. (Roivant Sciences, 2023).

73. B. E. Sands *et al.*, The Impact of Raising the Bar for Clinical Trials in Ulcerative Colitis. *J Crohns Colitis* **13**, 1217-1226 (2019).

74. S. Danese, E. Schabel, M. A. Ainsworth, L. Peyrin-Biroulet, Challenges and opportunities for IBD drug development: from early stage to regulatory approval. *Gut* **69**, 1157-1161 (2020).

75. M. C. Dubinsky, R. Collins, M. T. Abreu, I. O. f. t. S. o. I. B. D. (IOIBD), Challenges and Opportunities in IBD Clinical Trial Design. *Gastroenterology* **161**, 400-404 (2021).

76. T. Straatmijer *et al.*, Superior Effectiveness of Tofacitinib Compared to Vedolizumab in Anti-TNF-experienced Ulcerative Colitis Patients: A Nationwide Dutch Registry Study. *Clin Gastroenterol Hepatol* **21**, 182-191.e182 (2023).

77. R. S. Parra *et al.*, Long-term effectiveness and safety of ustekinumab in bio-naïve and bio-experienced anti-tumor necrosis factor patients with Crohn's disease: a real-world multicenter Brazilian study. *BMC Gastroenterol* **22**, 199 (2022).

78. H. K. Hyun *et al.*, Comparative effectiveness of second-line biological therapies for ulcerative colitis and Crohn's disease in patients with prior failure of anti-tumour necrosis factor treatment. *BMC Gastroenterol* **22**, 143 (2022).

79. B. Verstockt *et al.*, Influence of Drug Exposure on Vedolizumab-Induced Endoscopic Remission in Anti-Tumour Necrosis Factor [TNF] Naïve and Anti-TNF Exposed IBD Patients. *J Crohns Colitis* **14**, 332-341 (2020).

80. A. Buisson *et al.*, Real-world comparison of effectiveness between tofacitinib and vedolizumab in patients with ulcerative colitis exposed to at least one anti-TNF agent. *Aliment Pharmacol Ther* **57**, 676-688 (2023).

81. A. Aliper *et al.*, Prediction of Clinical Trials Outcomes Based on Target Choice and Clinical Trial Design with Multi-Modal Artificial Intelligence. *Clin Pharmacol Ther* **114**, 972-980 (2023).

82. W. A. Haynes, A. Tomczak, P. Khatri, Gene annotation bias impedes biomedical research. *Sci Rep* **8**, 1362 (2018).

83. J. A. Timmons, K. J. Szkop, I. J. Gallagher, Multiple sources of bias confound functional enrichment analysis of global -omics data. *Genome Biol* **16**, 186 (2015).

84. M. D. Young, M. J. Wakefield, G. K. Smyth, A. Oshlack, Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol* **11**, R14 (2010).

85. P. Gaudet, C. Dessimoz, Gene Ontology: Pitfalls, Biases, and Remedies. *Methods Mol Biol* **1446**, 189-205 (2017).

86. S. Mubeen, A. Tom Kodamullil, M. Hofmann-Apitius, D. Domingo-Fernández, On the influence of several factors on pathway enrichment analysis. *Brief Bioinform* **23**, (2022).

87. L. Liu, X. Dai, H. Wang, W. Song, J. Lu, A weighted multipath measurement based on gene ontology for estimating gene products similarity. *J Comput Biol* **21**, 964-974 (2014).

88. M. Milano, G. Agapito, P. H. Guzzi, M. Cannataro, in *2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. (2014), pp. 9-16.

89. F. Ren *et al.*, A small-molecule TNIK inhibitor targets fibrosis in preclinical and clinical models. *Nat Biotechnol*, (2024).

90. R. G. Langley *et al.*, Safety results from a pooled analysis of randomized, controlled phase II and III clinical trials and interim data from an open-label extension trial of the interleukin-12/23 monoclonal antibody, briakinumab, in moderate to severe psoriasis. *J Eur Acad Dermatol Venereol* **27**, 1252-1261 (2013).

91. T. Okada, T. Kajimoto, S. Jahangeer, S. Nakamura, Sphingosine kinase/sphingosine 1-phosphate signalling in central nervous system. *Cell Signal* **21**, 7-13 (2009).

92. A. Yazdi, H. Baharvand, M. Javan, Enhanced remyelination following lysolecithin-induced demyelination in mice under treatment with fingolimod (FTY720). *Neuroscience* **311**, 34-44 (2015).

93. B. G. Feagan *et al.*, Ozanimod as a novel oral small molecule therapy for the treatment of Crohn's disease: The YELLOWSTONE clinical trial program. *Contemp Clin Trials* **122**, 106958 (2022).

94. G. Alfonso Perez, R. Castillo, Gene Identification in Inflammatory Bowel Disease via a Machine Learning Approach. *Medicina (Kaunas)* **59**, (2023).

Acknowledgments: No editing services or AI-technologies were used in drafting any part of this manuscript.

Funding:

The Leona and Harry B. Helmsley Trust, Crohn's & Colitis Foundation, [036261] (PG, BSB)

National Institutes of Health grant R01AI155696, R01AI141630, UG3 TR003355, UH3 TR003355 (PG)

National Institutes of Health grant K23 DK123406 (BSB)

National Institutes of Health grant R01DK13593701A1, R03DK129631 (SiS)

American Association of Immunologists' (AAI) Intersect Fellowship Program for Computational Scientists and Immunologists (SaS)

AGA Research Foundation's AGA-Aman Armaan Ahmed Family Summer Undergraduate Research Fellowship-AGA2023-41-02 (EM)

Crohn's & Colitis Foundation award# 1036261 (EM)

Author contributions (CRediT):

Conceptualization: SaS, PG
Methodology & Resources: SaS, PG
Formal Analysis, Validation, Software: SaS
Consultation and domain expertise: BB, SiS.
Data Curation, Investigation & Visualization: SaS, EM, MM, PG
Funding acquisition: PG
Project administration: PG
Supervision: PG
Writing – original draft: SaS, PG
Writing – review & editing: SaS, EM, MM, SiS, BSB, PG

Competing interests: B.S.B receives research grants from Gilead, Merck, and has received consulting fees from Merck, Pfizer and Bristol Myers Squibb. SiS receives research grants from Pfizer. The authors declare no financial associations or other conflicts of interest. The remaining authors declare that they have no competing interests.

Data and materials availability: All data are available in the main text or the supplementary materials. All the publicly available datasets that were used in this work are listed in Supplemental Dataset 1. All the codes used in this work will be made public at the time of acceptance and available through password-protected links on Github during peer-review.

Supplementary Materials

Materials and Methods

Figs. S1 to S3

References (*1-94*)

Datasheet (Excel) S1 to S3

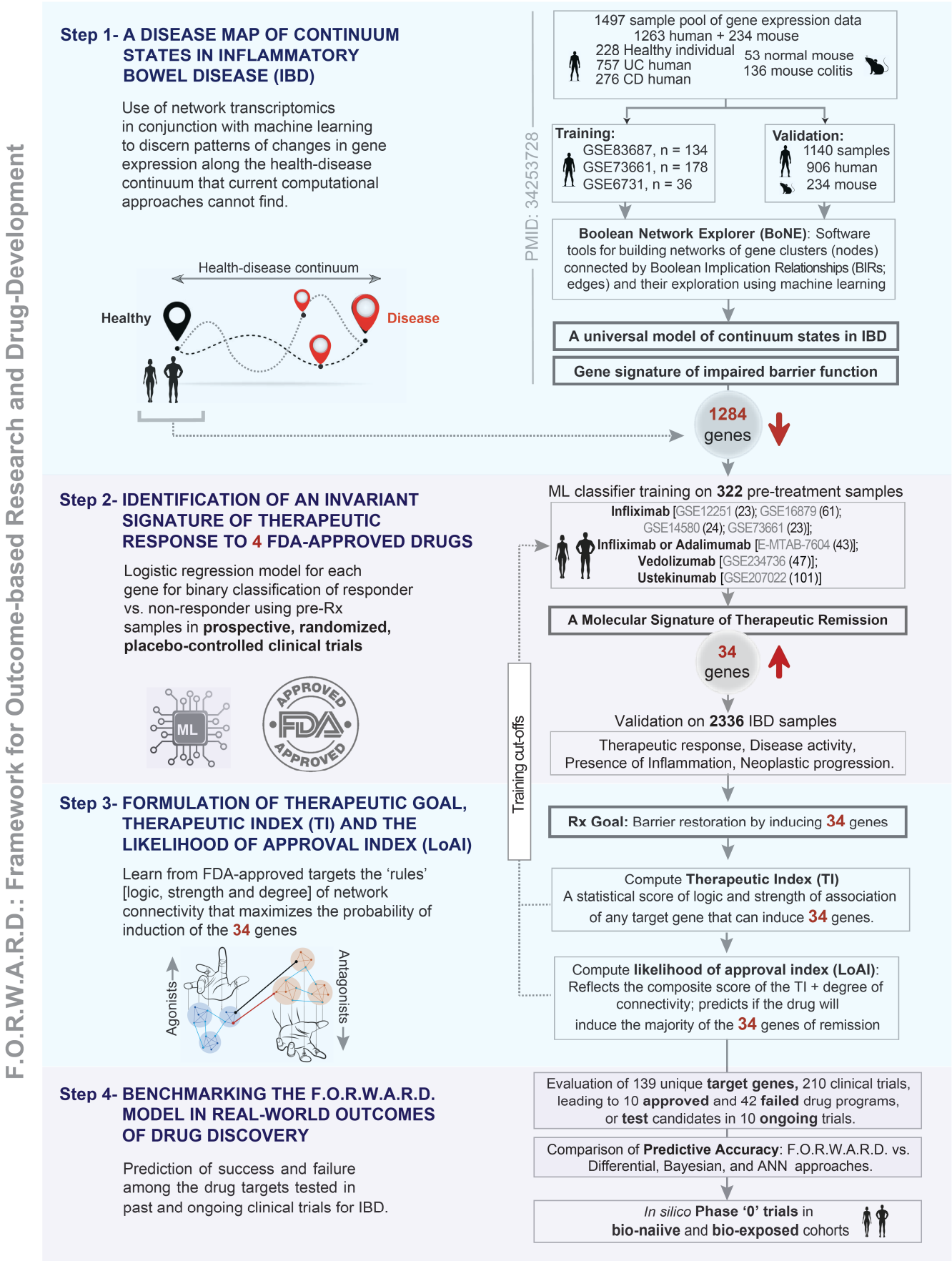

Figure 1. ***Study motivation, rationale and design.***

The ***four steps*** of a workflow that led to the development of a framework for outcome-based drug target prioritization (F.O.R.W.A.R.D) are presented. ***Step 1****:* A published Boolean implication network-based computational model of disease continuum states in IBD (*20*), which serves as the premise of this study. The model was created and rigorously validated on transcriptomics databases containing overall 1497 gene–expression data (1263 human and 234 mouse samples). ***Step 2****:* Use of machine learning (ML) classifier to identify a 34 gene molecular signature of barrier restoration that is shared among all responders, regardless of therapeutic modality. Outcome-annotated datasets from real-world randomized, placebo-controlled prospective clinical trials that were used for training and validation are shown. ***Step 3****:* Formulation of a Boolean implication-based connectivity index based on outcome-specific learning from real-world drug trials to standardize the model. Here, statistical learning is used to determine the logic and strength (TI) and degree (LoAI) of connectivity of drug's target genes to the 34 genes. A LoAI >/= 0.50 was determined as likely to pass regulatory approval. ***Step 4***: Benchmarking of F.O.R.W.A.R.D was done by two approaches: (i) test its predictive accuracy on drug development programs with known outcomes for the presence or lack of efficacy; (ii) test its ability to inform trial design through *in silico* trials on custom cohorts.

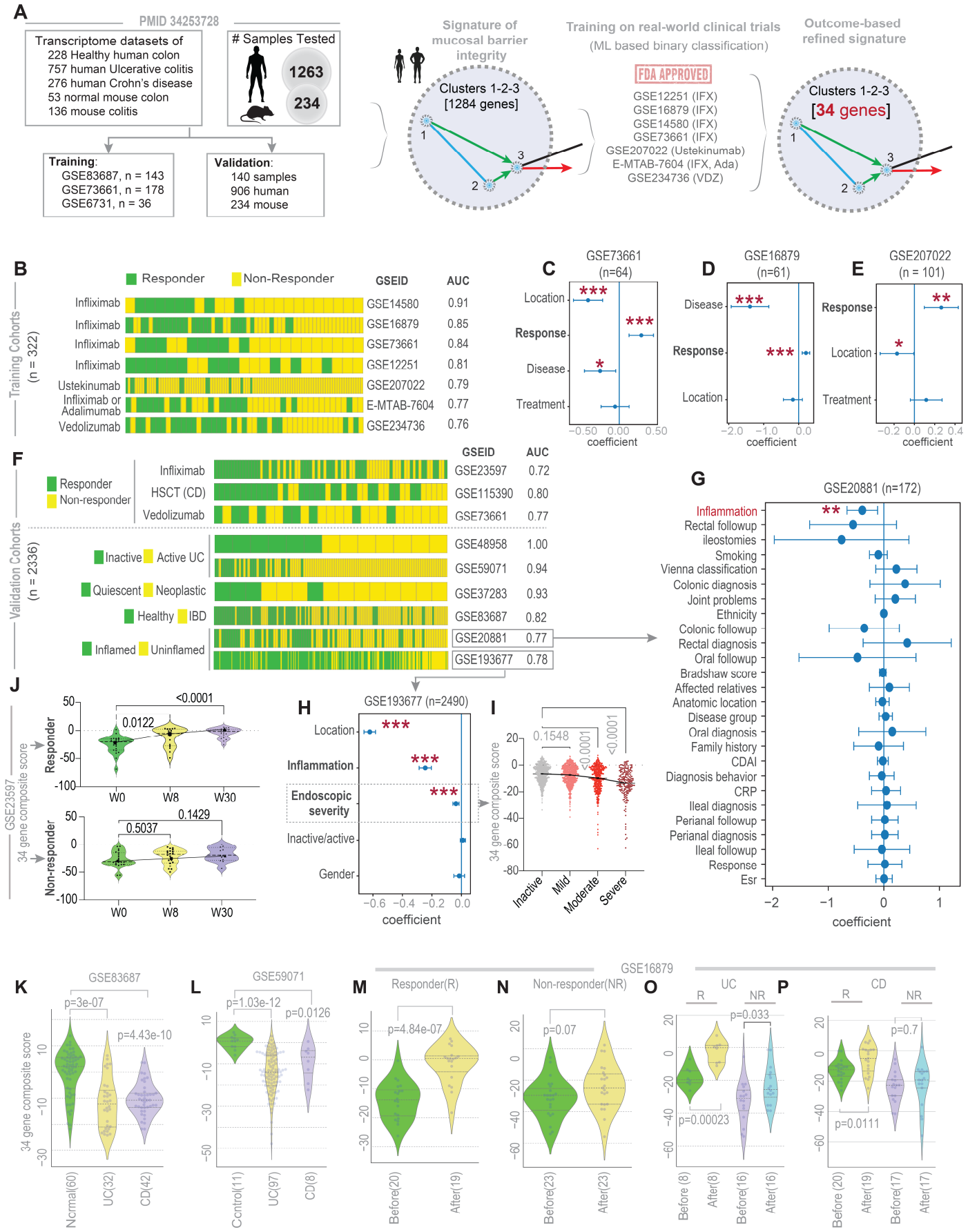

Figure 2. ***Identification of a molecular signature of therapeutic remission and derivation of a therapeutic goal.***

A. The computational workflow and the number of unique IBD samples that led to the identification of a 34 gene molecular signature of therapeutic remission. *Left*: A total of 1284 genes (in gene clusters# 1, 2, and 3 of a previously validated Boolean IBD network(*20*)) that were identified using machine learning (ML) as the invariant signature of mucosal barrier integrity in all samples across species. *Middle*: The signature of barrier integrity is subsequently trained using ML-based classifier on pre-treatment samples in 7 cohorts reflecting real world randomized, placebo-controlled prospective clinical trials on 4 different FDA-approved drugs. *Right*: A subset of 34 of the 1284 genes emerged as a robust classifier of responders from non-responders regardless of cohort heterogeneity or the drug class tested. **B**. Bar plots display the ability of the 34 gene signature to classify responder vs non-responder in the training datasets. Performance is measured by computing ROC-AUC. The rank order of different samples on the bar plot is based on the composite scores of 34 genes, arranged from left (high score) to right (low score). HSCT, Hematopoietic stem cell transplantation. **C-E**. Multivariate analyses on the training cohorts (in B) that reported >/= 3 variables (including 'response' status) using the composite score of 34 genes as a linear combination of all variables. The coefficient of each variable (at the center) with 95% CI (as error bars) and the p-values are illustrated. The p-value for each tests the null hypothesis that the coefficient is equal to zero (no effect). **F**. Bar plots show the ability of the composite expression score of the 34 genes in rank-ordering IBD disease groups, as in B, in various discovery cohorts. The disease groups include: (*from top to bottom*) response status in 3 additional treatment cohorts (Infliximab, GSE23597; Vedolizumab, GSE73661; and hematopoietic stem cells therapy, GSE115390), disease activity, evidence for neoplastic progression, presence or absence of disease, and presence or absence inflammation. **G-H**. Multivariate analyses show the composite score of 34 gene as a linear combination of all other variables. Endoscopic severity in panel H is used as a continuous variable. The p-value for each term tests the null hypothesis that the coefficient is equal to zero (no effect). p-values are as follows: * = $p \leq 0.05$, ** = $p \leq 0.01$, and *** = $p \leq 0.001$. **I**. Scatter plots display the composite score of 34 genes in samples annotated for varying endoscopic severity. The p values, calculated by Welch's t-test are displayed. **J-P**. Violin plots display the composite score of the levels of expression the 34 genes in patients receiving biologics, categorized based on treatment duration (J), disease type (K-L), timing of biopsy (before vs after biologics; M-N) and response status by disease subtype (UC vs CD; O-P). p-values were calculated by Welch's t-test. See also **Fig S1A-D** for functional enrichment analyses on the 34 gene signature. A catalog of all datasets and patient characteristics used in this work can be found in **Supplemental Datasheet 1**.

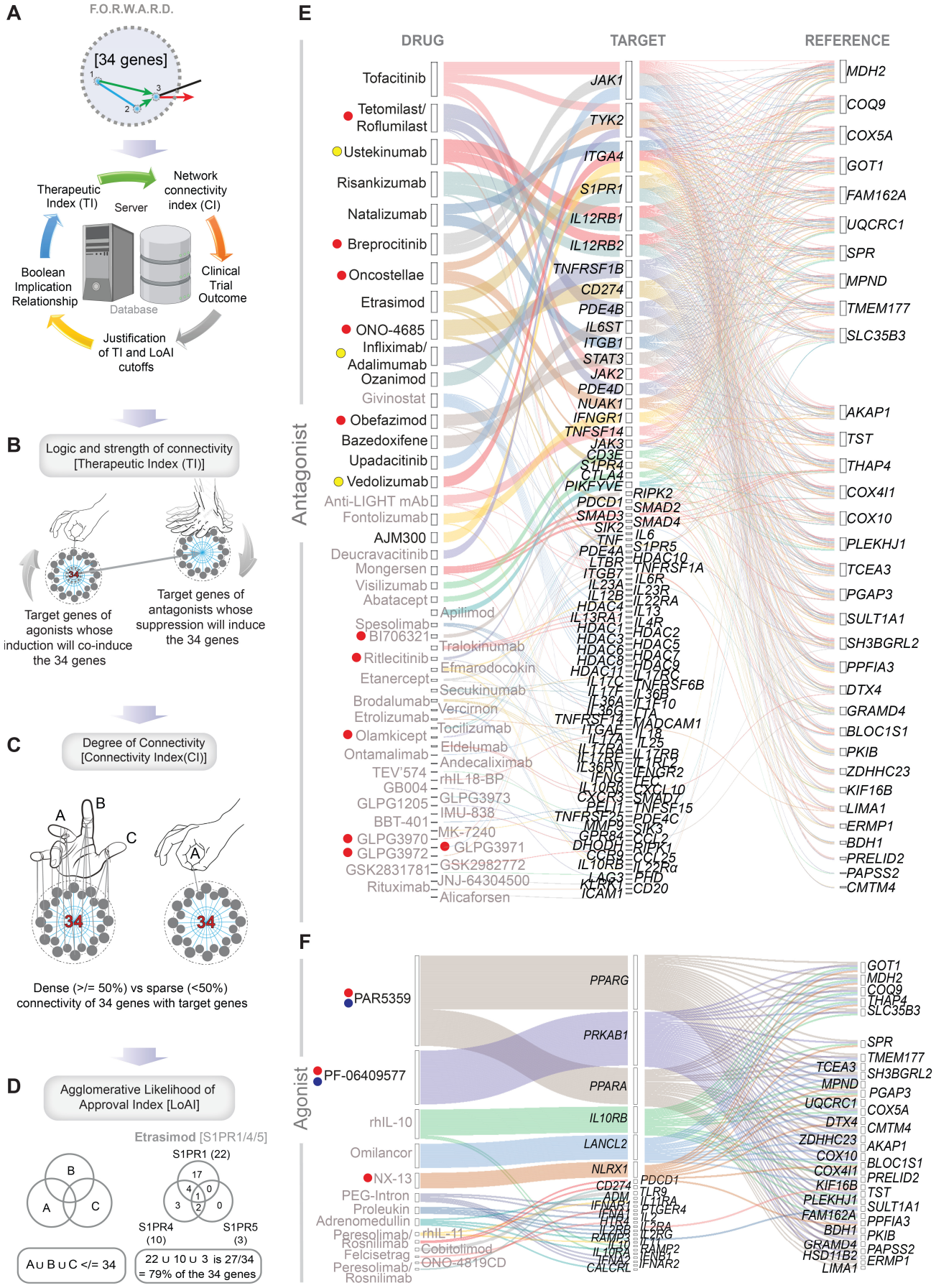

Figure 3. Connectivity of drug target gene(s) to the 34 gene molecular signature of therapeutic remission.

A. Schematics illustrates the various statistical derivatives within the computational framework of F.O.R.W.A.R.D that enable objective scoring of the logic, strength and degree of connectivity of any drug's target gene(s) with the 34 reference genes that track the likelihood of therapeutic remission. **B.** The logic and strength of connectivity is assessed by calculating 'therapeutic index,' (see *Methods* and **Fig S2**). Drugs acting as *ago*nists are expected to co-induce the 34 reference genes via equivalent Boolean Implication Relationships (BIRs), 'T equivalent R' and 'T high=> R high' (*left*), whereas drugs acting as *anta*gonists are expected to induce the 34 reference genes via opposing BIRs, 'T opposite R' and 'T low=> R high' (*right*). **C.** Connectivity index (CI) is a number between 0 and 1, which reflects the fraction of the 34 reference genes to which target gene relates to a TI <0.075. Connectivity is deemed as sparse vs dense using a >/= 50% as cutoff, based on the simple majority protocol. **D**. Schematic (*left*) showcases how the agglomerative likelihood of approval index (LoAI) for a generic drug is derived based on the union summation of its target gene's connectivity with the 34 gene signature of remission. Schematic (*right*) cites the process of LoAI estimation for *Etrasimod*, which targets proteins encoded by 3 distinct genes. The CI for each target is indicated in parentheses. **E-F.** Alluvial plots for antagonists (E) and agonists (F) representing the connectivity stream between drug names (*left column*), major score-driving target gene (*middle column*) and the reference genes within the 34 gene signature of therapeutic remission (*right column*). A comprehensive list of target genes, TI, CI, and LoAI values, and their connectivity with reference genes are shown in **Supplemental datasheet 2**. *Black* and *grey* fonts indicate drugs with an LoAI of at least >/= 0.5 (i.e., achieved simple majority) or below (fails to reach majority), respectively. Red dots (in E-F) mark drugs in ongoing trials where outcomes are unknown currently. Yellow dots (E) mark the 4 FDA-approved drugs used for model training. Blue dots (F) mark promising preclinical targets identified using Boolean Implication network approach.

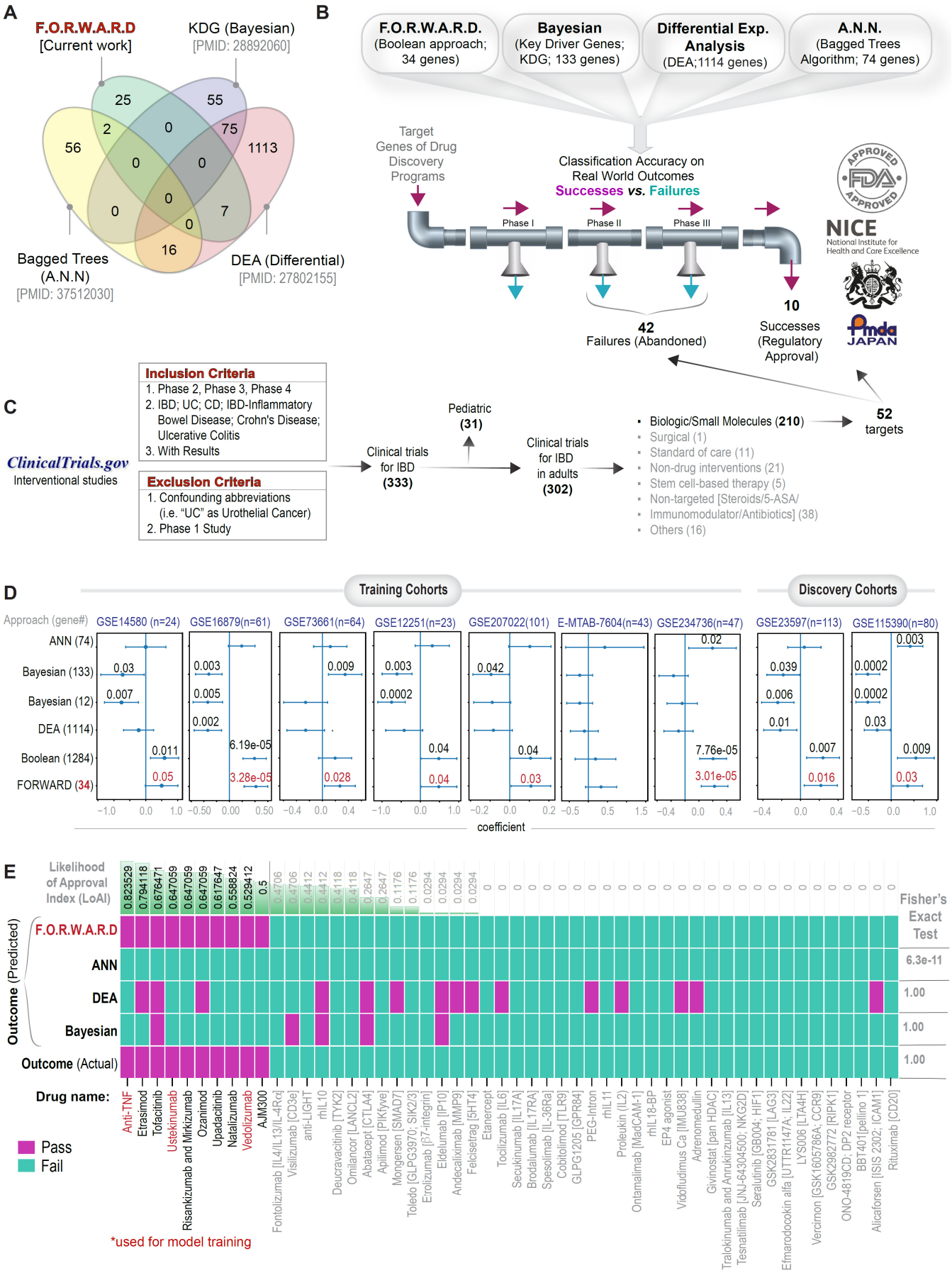

Figure 4. Predictive accuracy of *F.O.R.W.A.R.D* on drug development programs with known outcomes.

A. Venn diagram illustrating the overlap of genes identified by various methods: the Bayesian KDG algorithm(*42*), Differential expression analysis (DEA) between healthy individuals and IBD patients(*56*), the top-most performing artificial neural network (A.N.N) model, i.e., bagged tree algorithm (*94*), and the 34 genes used in F.O.R.W.A.R.D (current work). **B-C.** Schematic (B) and flowchart (C) illustrate the process by which drug programs were chosen for their inclusion in benchmarking studies. A catalog of all drug programs included in this analysis can be found in **Supplemental Datasheet 2** and gene signatures in **Supplemental Datasheet 3**. **D**. Univariate analyses presenting the composite scores of various gene sets in A, and their related signatures, i.e., the 12 top-ranked KDGs that is a subset of the 133 Bayesian KDGs(*42*), and the superset of 1284 genes, from which the 34 genes were derived for the F.O.R.W.A.R.D algorithm (**Fig 2A**, current work). Results are displayed as a linear combination of all the composite scores with the response status of patients across all training and validation cohorts. The coefficient of each variable (centered) is displayed with 95% confidence intervals (error bars) and the numbers display p-values. The p-value for each tests the null hypothesis that the coefficient is equal to zero (no effect). **E.** A binary matrix heatmap shows the pass (magenta) and fail (teal) prediction on 52targets representing 210 clinical trials based on F.O.R.W.A.R.D and 3 conventional methods that are currently used to prioritize target genes. The three drugs that were used for training F.O.R.W.A.R.D are highlighted in red. The actual outcome is annotated in the bottom row. The Likelihood of approval (LoAI) score derived from the F.O.R.W.A.R.D framework is displayed across the top row. Each drug's target genes, and their TI, CI and LoAI are presented in **Supplemental datasheet 2.**

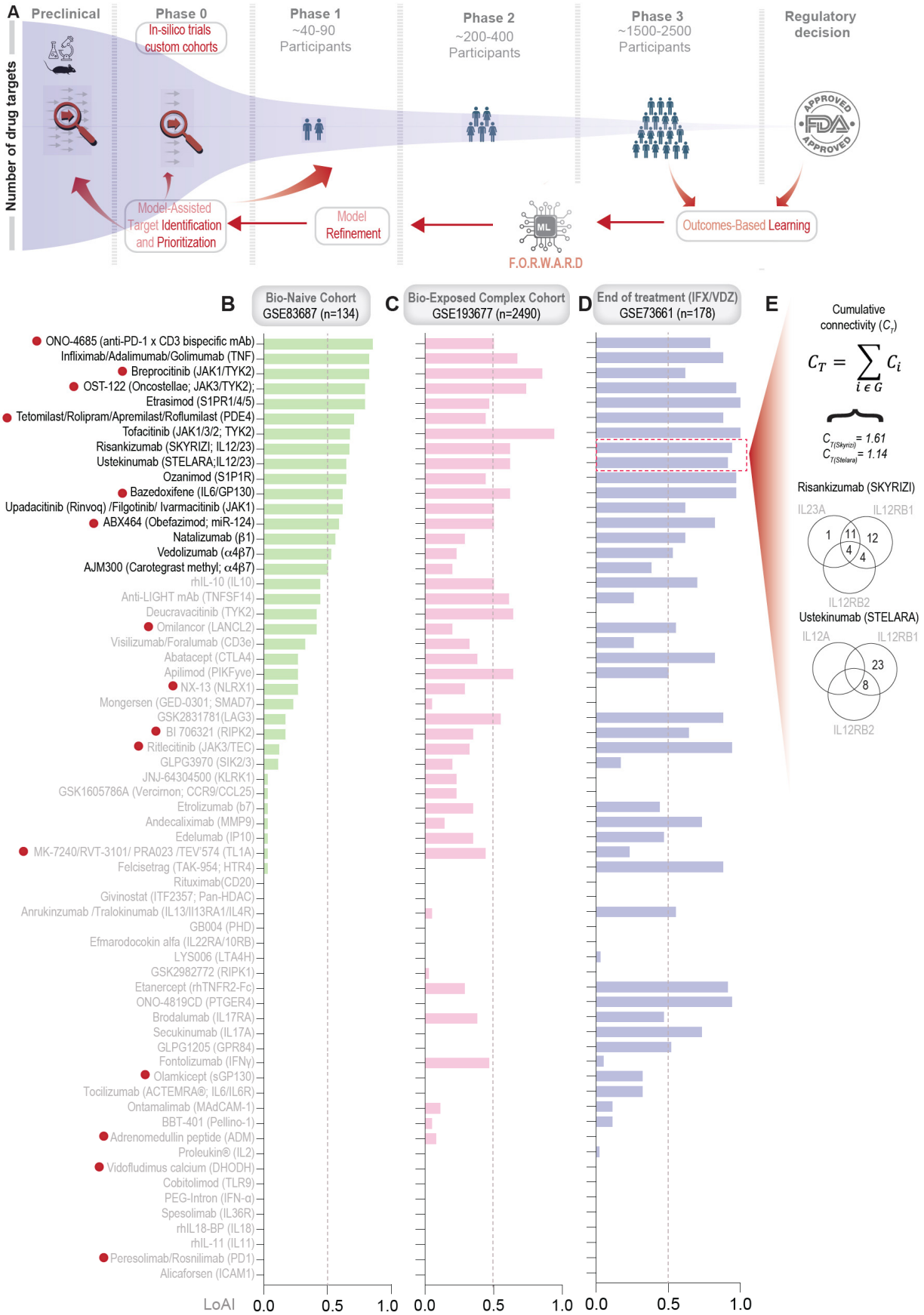

Figure 5. Integration of F.O.R.W.A.R.D into the current drug discovery process through *in silico* trials.

A. Schematic shows how the F.O.R.W.A.R.D framework could be seamlessly integrated into both the early and late stages of the drug development process (*red* arrows) and the potential impacts of such integration (*red text* in boxes). **B-D**. Bar plots show the likelihood of approval index (LoAI) of drugs and their major target genes (listed in parenthesis) in three custom cohorts reflecting important clinically directive characteristics considered during trial design and recruitment. Red dots next to a drug's name implies that these are ongoing trials. IFX, infliximab; VDZ, vedolizumab. **E.** Cumulative connectivity, which is an estimate of total overlapping connectivity of multi-target drugs with the 34-gene signature of response, is shown for risankizumab, mirikizumab and ustekinumab, as determined in a cohort of patients exposed to anti-TNF or vedolizumab.