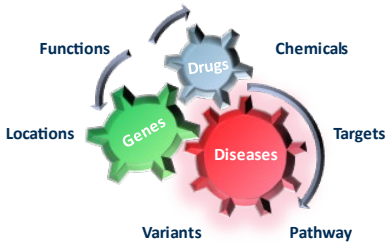
THEMATIC

$7-9bn 2030E

AI Drug Discovery Market



Biomedical relationships in drug discovery (Sources: ACF Equity Research Graphics; QIAGEN).

Thursday, 22 September 2022

Peer Median 52wk Δ	-56%
AI Drug Discovery Mrkt Est.$(m) 2030E	$10,811
Media Mrkt Est.$(m) 2027E	$3,200
Currency	USD
Business Activity	
AI Drug Discovery & Development	
Key Metrics $(m) 2022E	
Av. Cost Drug Est.	$3,200
Median Cost Drug Est.	$1,000
AI saving per Drug Est.	$115-350
Key Ratios	
2022E Vintage Est.	0.80
FX Rate USD/USD	1.00

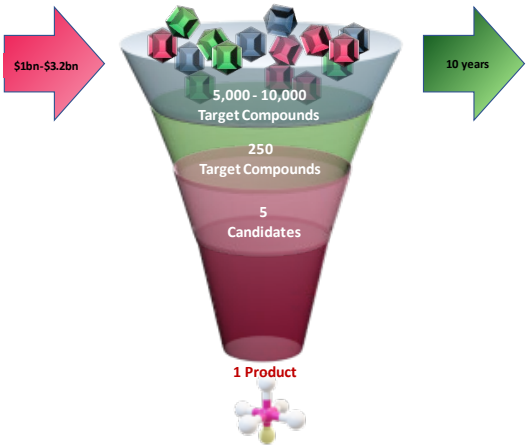
Healthcare Sector Research Thematic Analyst Team
+44 20 7419 7928
healthcare@acfequityresearch.com

The Future of Drug Discovery

Drug discovery and development requires colossal upfront investment. The process typically takes over a decade. A large number of product candidates fail to prove efficacy or safety during preclinical and clinical research phases or fail to pass regulatory scrutiny. According to Deloitte, the average cost of drug discovery and development increased over 80% between 2010 and 2018 from $1.188bn to $2.168bn, CAGR 7.8%, per successfully launched drug. By inflation only, we would expect costs to have risen to ~$1.5bn by 2018 and ~$1.8bn by YE22E, assuming an average inflation rate of 2.5% p.a. from YE09A to YE21A and 13% in YE22E. Inflation is not the key driver, and so we estimate the average cost from discovery to drug launch YE22E could exceed ~$3.2bn. The median cost may be somewhat lower at around $1bn.

- $1bn is our YE22E median estimated new drug launch cost;
- $3.2bn is our YE22E average estimated new drug launch cost;
- 25%-33% - our cost saving estimate for preclinical AI drug discovery;
- Cash savings potential estimated $115m to $1bn per drug launched.

Sources: ACF Equity Research Graphics.



Market – AI Drug Discovery Opportunity

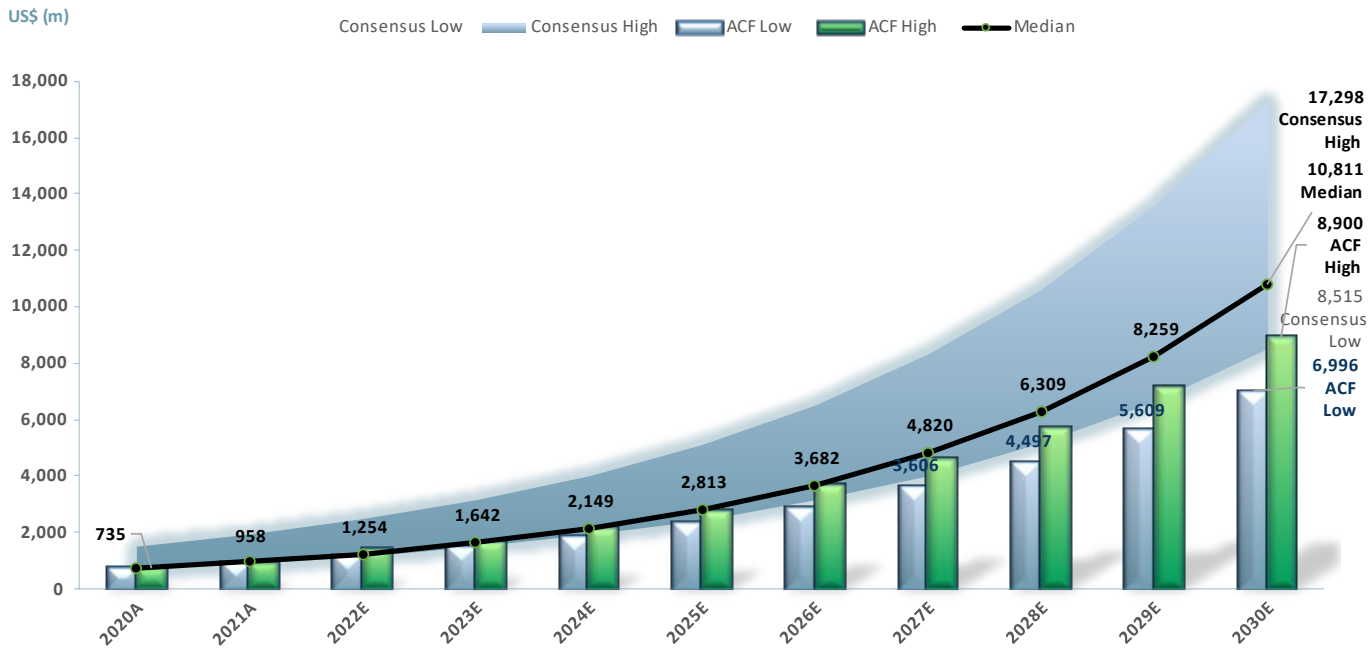
Our forecasts for the value of the AI drug discovery market **2030E value, range between US$ 6,996m and $8,900m; low case CAGR 24.7%, high case CAGR 28.1%**.

In the exhibit below our forecasts are at the low end of the consensus range and our low case is below the consensus range. **The market consensus spread of $8,783m, (high $17,298m, low $8,515m, median $10,811m, high CAGR 37.6%, low CAGR 27.7%)** is composed of five other recognized and respected market forecast providers in the AI drug discovery market. In certain cases, we have normalized the consensus forecasts in order to extend them to 2030E.

Whichever way we approach this market forecast the observations are broadly the same – the view on future value has a very wide-spread, with the top end market consensus forecast being over 100% higher than the bottom end. Each forecast differs in the specifics of how it defines the AI market for drug discovery.

We applied a statistical analysis to generate our forecasts. We have overlayed various factors including the effect of Covid, higher interest rates, and current market sentiment – i.e. the ability of companies to raise money and so invest in the AI drug development technologies or start new drug development projects. Other drivers are rising costs of drug development, time, wastage, patient outcomes, M&A and or alliance forming.

Exhibit 1: **AI drug discovery market forecasts**



Sources: ACF Equity Research; GrandView Research; Precedence Research; Emersion Insights; Facts and Factors; Novaone Advisors.

Drug discovery is an extremely complex process subject to significant error probabilities. The difficulty in finding new drugs and cell growth regulating protein targets lies in the high number of variables associated with biological processes. For example, one protein growth regulator target may affect multiple downstream protein regulator targets.

This waterfall effect quickly leads to an 'exponential' number of potential impacts. These interactions eventually affect upstream pathways, such as differentiation, migration and apoptosis, themselves affecting the original protein target and or affecting other unintended biological pathways.

Current drug discovery and development approaches are error prone because they rely on historical knowledge of biological signaling pathways combined with investigator/scientist intuition when searching for new solutions and new pathways.

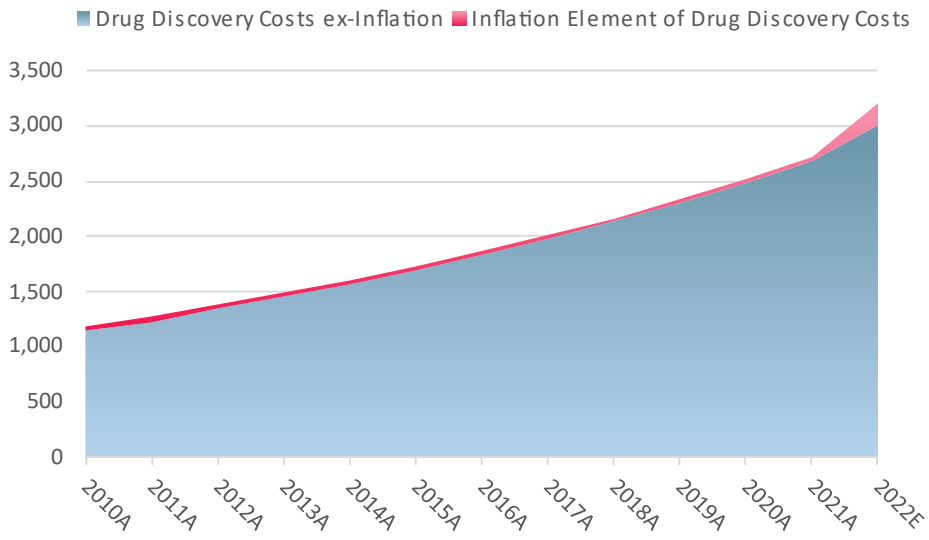
Implies thousands of years of analysis.

The level of complexity and difficulty can also be considered this way: Given a hypothetical multivariate analysis approach to understanding the effects of different drug effects on different target proteins in different cells, it was estimated that the entire experimental space would require about 10 billion plates for analysis (there are hundreds of wells on each plate).

Hormesis is a biological characteristic in which exposure to toxins and other stressors is favourable. There is often a biphasic of triphasic response to increasing exposure to toxins.

The implication of 10 billion plates with hundreds of wells is thousands of years of analysis. Adding variables such as drug combinations, cell combinations and or drug dose (which can be critical since many drugs may have hormetic dose responses), increases complexity and so the difficulty of the analysis.

Exhibit 2: **Inflation impact on drug discovery costs 2010A-2022E**



Sources: ACF Equity Research Estimates; Deloitte.

8 steps to delivering a new drug to market.

A closer look at the opportunity - One cited reason for high drug prices is the total cost and high failure rates for bringing drugs to market through all of the stages of development, which we break into 8 steps here (though it could be counted as more):

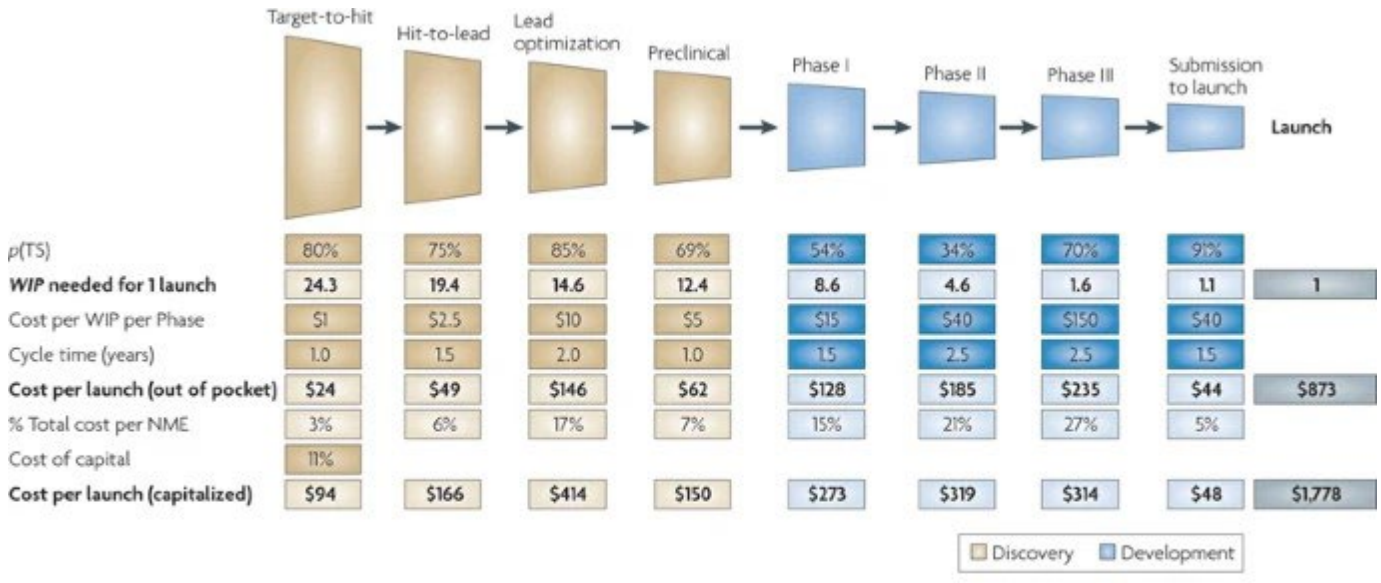
Steps 1-4 have an average cost of $600-700m or ~33% of per successfully launched drug.

1. Target-to-hit;
2. Hit-to-lead;
3. Lead optimization;
4. Preclinical;
5.-7. Clinical phase I, II and III;
8. Regulatory submission.

According to various estimates, the cost of preclinical research, that is, finding a target, finding a drug for that target, optimizing the drug, and testing it in animals and ex vivo, can cost $600-700m per launched (successful) drug, or ~33% of the entire cost of a new drug.

Improvements early on in the development process can lower overall costs by reducing failure rates (gaining better insights that translate into clinical trials such as potential hepatotoxicity, biomarkers for the proper patients, etc.), and making the process more efficient (using AI to quickly identify the right drugs for certain biological targets using approaches like the active learning cycle driven by AI.

Exhibit 3: **Drug discovery & development segmented costs estimate**



Sources: Nature Reviews Drug Discovery; Paul, S., Mytelka, D., Dunwiddie, C. et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nat Rev Drug Discov 9, 203–214 (2010).

AI Drug Development – Costs Transformation

Drug discovery is the first phase of the value chain.

Artificial Intelligence is transforming Biopharma R&D drug discovery and development. Drug discovery is the first phase of the value chain that identifies new candidates or therapies for treating human disease.

Drug discovery is the initial stage of R&D.

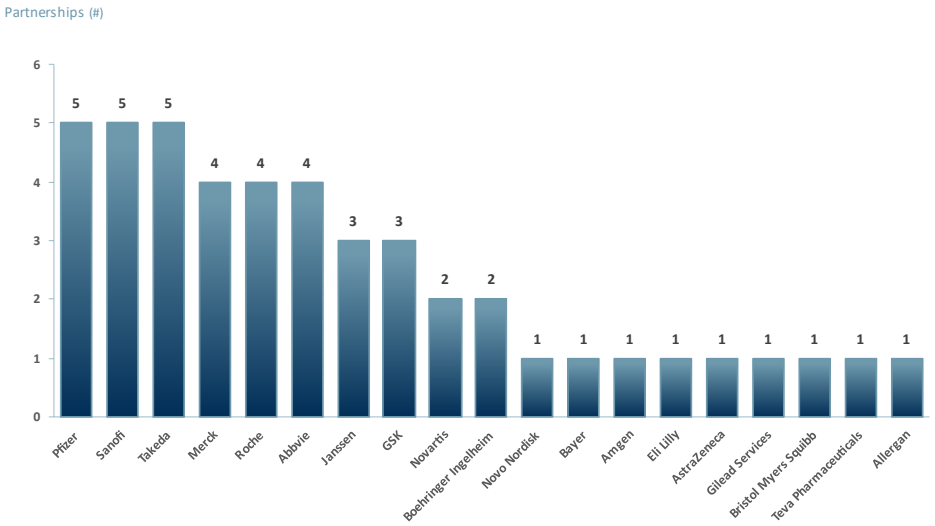
Drug discovery is also the initial stage of R&D that involves the identification and optimization of potential new drugs and preclinical in vivo validation through cell assays and animal models. After the development stage, successful candidates that meet the regulatory requirements move into the clinical trial phase where drugs are tested on humans, typically in three phases, PI, PII, and PIII.

Drug discovery is high-risk.

AI drug development has potential to improve risk, time and cost profiles of drug discovery.

Drug discovery is a high-risk, high capital, high barrier to entry, enterprise. AI drug development has potential to improve both the risk, time and cost profile of drug discovery. The potential of AI in drug discovery is not lost on big pharma (see exhibit below).

Exhibit 4: **Big pharma AI partnerships disclosed to market 2021A**



Source: Market Future Research.

Preclinical research costs est. at $600-700m per launched (successful) drug.

~33% of the entire cost of a new drug.

Cost of preclinical research - Market estimates suggest the cost of preclinical research, which is finding a target, finding a drug for that target, optimizing the drug, and testing it in animals and ex vivo, can cost **$600-700m per launched (successful) drug, or ~33% of the entire cost of a new drug.** Improvements early on in the development process can lower overall costs by reducing failure rates. AI contributes to reducing failure rates by quickly identify the right drugs for certain biological targets.

In preclinical research alone, unvalidated media and reagents cost pharmaceutical companies over $28bn in 2015.

Waste - Preclinical research using unvalidated media and reagents cost pharmaceutical companies over $28bn in 2015, due to experimental irreproducibility (bad studies) rates exceeding 50%, in preclinical research alone. Thus, it is implied that there is further waste in clinical research as well.

The average probability of moving from PII to PIII is <33% and from PIII is little above 35%

Failure rates in phase II and III (PII and PIII) oncology research tend to be comparatively high when compared to phase I (PI) failure. This suggests that PI results lead to false corporate and investor perceptions of project de-risking. PI trials are often dose-escalation and safety analyses.

We infer that huge value generation opportunities exist if AI drug discovery delivers significant failure rate improvements.

Some research suggests that the **average probability of moving from PII to PIII is <33%** and from PIII is little above 35% (see exhibit below). This suggests that the preclinical models used to determine efficacy of a drug correlate very poorly with the heterogeneity of human disease. We infer that huge value generation opportunities exist if AI drug discovery delivers significant failure rate improvements.

>50% of biomedical researchers "do not bother to verify the identity of their cell lines."

The future value generator potential of companies working to fine-tune and control the quality of future preclinical research is difficult to overestimate.

Imposter cell lines – We have come across academic research that suggests that ~33% of all cell lines are imposters (the wrong cell line—i.e., melanoma instead of thyroid cancer). In the scientific journal, Nature, it was suggested that >50% of biomedical researchers "do not bother to verify the identity of their cell lines." The cost of flawed preclinical research possibly runs into billions of USD. Unfortunately, flawed research programs also influence and 'inform' future research. In our view, the future value generator potential of companies working to fine-tune and control the quality of future preclinical research is difficult to overestimate.

AI might speed drug discovery by 40%, while cutting costs by 60%.

AI in drug discovery could be extremely useful in increasing the chances of success in mid-and-later stages of research (PII and PIII). More accurate preclinical research could increase the chances of approval success. Our tentative estimation is that AI might speed drug discovery by 40%, while cutting costs by 60%.

Exhibit 5: **Success probability by trial phase and therapeutic market**

Therapeutic Area	P1 to P2	P2 to P3	3 to Approval	Overall
Oncology	57.6	32.7	35.5	3.4
Metabolic / Endocrinology	76.2	59.7	51.6	19.6
Cardiovascular	73.3	65.7	62.2	25.5
Central Nervous System	73.2	51.9	51.1	15.0
Autoimmune / Inflammation	69.8	45.7	63.7	15.1
Genitourinary	68.7	57.1	66.5	21.6
Infectious Disease	70.1	58.3	75.3	25.2
Ophthalmology	87.1	60.7	74.9	32.6
Vaccine (Infectious Disease)	76.8	58.2	85.4	33.4
Overall	66.4	48.6	59.0	13.8
Overall (Excluding Oncology)	73.0	55.7	63.6	20.9

Sources: ACF Equity Research; Berzow, A (2020, June 12) Clinical Trial Success Rates by Phase and Therapeutic Area.

AI driven active learning cycle of drug discovery – a better approach?

We infer therefore, that a better approach to drug discovery and development is the use of AI driven guided iterative solutions, sometimes referred to as the active learning cycle of drug discovery.

Irreproducibility factor comes from poor design experimental design or incorrect conclusions and inferences and renders experiments flawed.

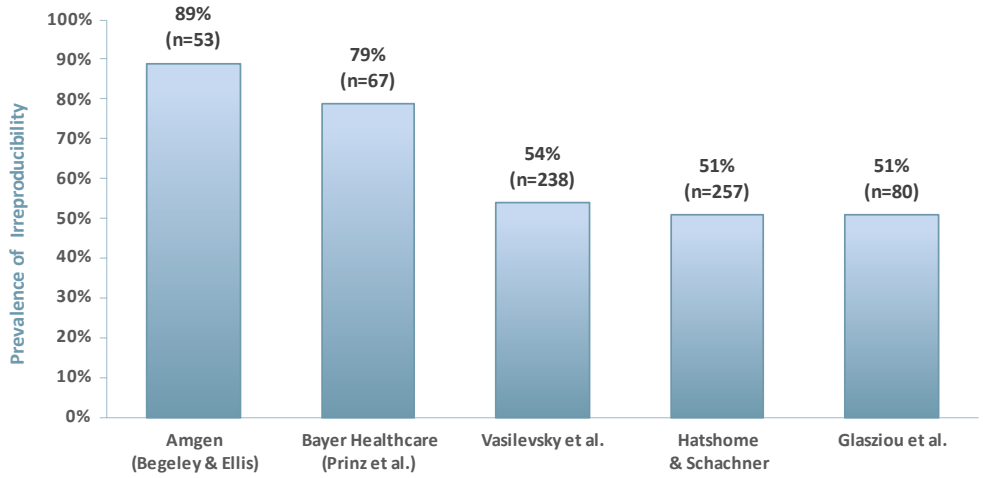
The problem of irreproducibility is adding to short and long run costs. Experiments are the backbone of preclinical research and innovation. Experiments are used to formally test hypotheses and gain more insight into biomolecular interactions and pathways. An essential factor in experimental research is reproducibility. If an experimental outcome cannot be reproduced when the experiment is repeated either in the original lab and or in any other lab, it can lead to significant wastage.

Waste from irreproducibility in preclinical research using unvalidated media and reagents costs the pharmaceutical sector companies over $28bn YE15A.

This so-called irreproducibility factor comes from poor experimental design or incorrect conclusions and inferences and renders experiments flawed in part or in whole.

According to a study on the economics of reproducibility of preclinical research, Freedman et al. estimated that waste from irreproducibility in preclinical research using unvalidated media and reagents costs the pharmaceutical sector companies over $28bn YE15A. Freedman's paper suggests this loss is due to irreproducibility rates in experiments exceeding 50% in preclinical research (it excludes losses beyond preclinical research).

Exhibit 6: **Prevalence of irreproducibility in drug R&D YE15A**



Sources: Begley and Ellis [6], Prinz et al. [7], Vasilevsky [8], Hartshorne and Schachner [5], and Glasziou et al. [9]; Freedman, L. P et al. (2015). The Economics of Reproducibility in Preclinical Research. PLOS Biology, 13(6).

Counterintuitively, these data do not mean that flawed studies have no utility, and Freedman et al. cite that the actual dollar amount wasted could be higher or lower based on the assumptions they themselves have made.

We infer there is a lot of waste in research regardless of the exact amount, and specifically in preclinical research.

30% of all cell lines are imposters.

We infer from the Freedman paper and anecdotal information that there is a lot of waste in research regardless of the exact amount, and specifically in preclinical research, due to issues such as unvalidated or flawed reagents/materials, study design, data analysis, reporting, and laboratory protocols.

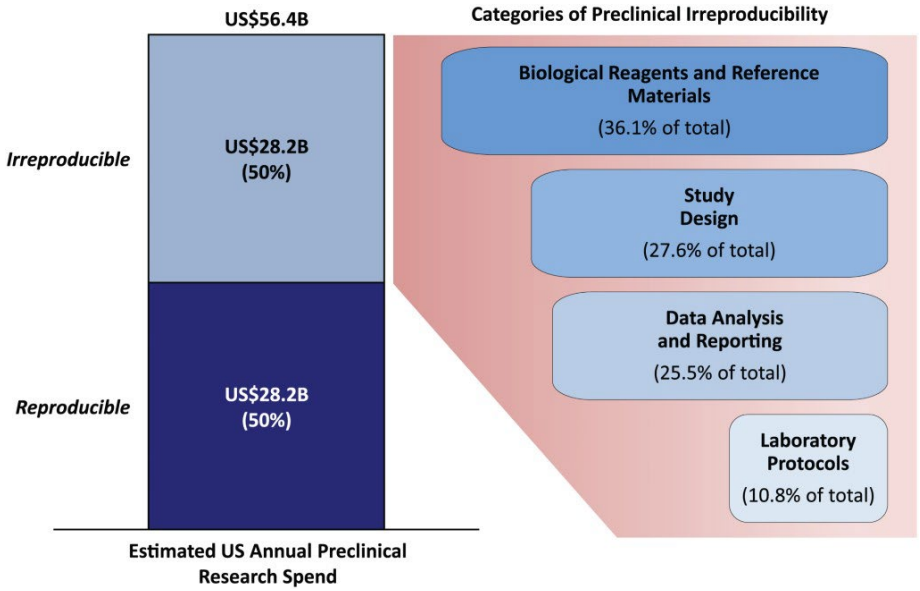
From the research conclusions we have outlined above, which support anecdotal information, it seems uncontroversial to infer that billions of investor dollars are wasted every year on flawed preclinical research.

Each study replication requires between 3 and 24 months and between $500k to $2m investment, flawed research driving further research is a significant ROI investor issue.

Flawed preclinical research also goes on to inform future research (with, for example, an additional 10,000 citations each year on false cell lines). As each study replication requires between 3 and 24 months and between $500k to $2m investment, flawed research driving further research is a significant ROI investor issue.

Axiomatically, the value generation from innovative approaches such as the AI driven active learning cycle used to fine-tune and control the quality of future preclinical research is difficult to overestimate.

Exhibit 7: **Wastage in pre-clinical R&D drug spend**



Sources: Freedman, L. P et al. (2015). The Economics of Reproducibility in Preclinical Research. *PLOS Biology, 13*(6).

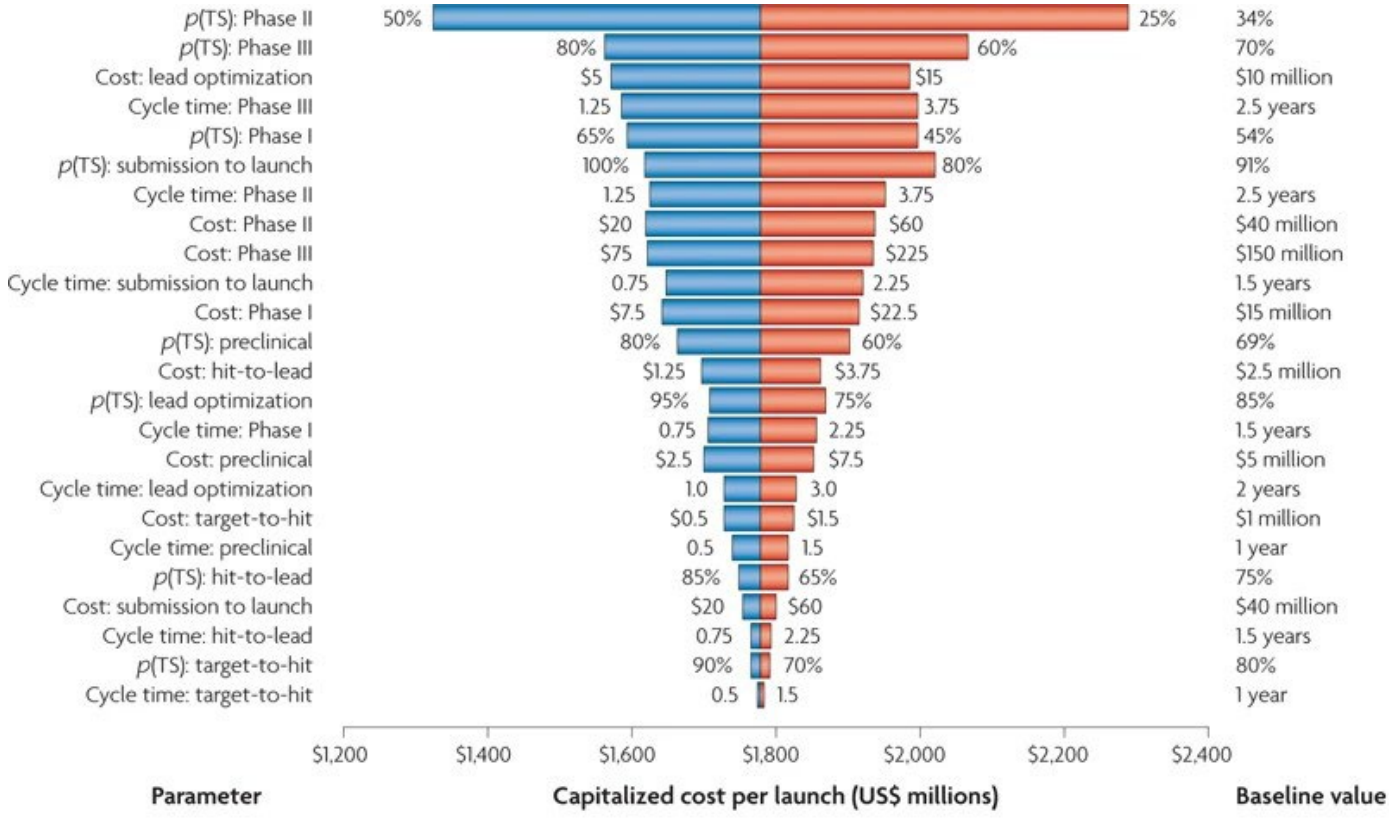
The cost estimates we provide here do not account for clinical trial failures in studies downstream, which might be avoided by better study designs that could come out of AI driven approaches.

Predictive Oncology estimates that its CoRE/PeDAL AI platform can speed discovery up by 40% while cutting costs by 60%.

Companies deploying AI driven drug discovery solutions believe they can deliver substantive savings. Predictive Oncology (NasdaqCM : POAI) estimates that its CoRE/PeDAL AI platform can speed discovery up by 40% while cutting costs by 60%. If such claims are only 60% true in the long run, there remains an implied ~25% time saving and ~35% cost saving on discovery per launched drug. On a $3bn drug-to-launch campaign assuming 33% of the current average cost is associated with discovery, our assumption suggests an average saving of $350m. If the median of the average costs for each successful drug launch is $1bn that suggests a saving of $115m.

We estimate AI drug discovery average savings could range from $115m - $350m per successful drug launch.

Exhibit 8: **Risks, costs and time of each phase of drug R&D**



Sources: Nature Reviews Drug Discovery; Sources: Paul, S., Mytelka, D., Dunwiddie, C. et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nat Rev Drug Discov 9, 203–214 (2010).

In the tornado exhibit above this "parametric sensitivity analysis" is created from an R&D model that calculates the capitalized cost per launch based on assumptions for the probability of technical success (p(TS)), cost and cycle time, by phase. When baseline values for each parameter are applied, the model calculates a capitalized cost per launch of US$1,778m - the spine of the sensitivity analysis.

The analysis above varies each of the parameters individually to a high and a low value (while holding all other parameters constant at their base value) and calculates a capitalized cost per launch based on those new values for that varied parameter.

In the analysis, the values of the parameters are varied from 50% lower and 50% higher relative to the baseline value for cost and cycle time and plus or minus 10 percentage points for the probability of technical success p(TS).

Once cost per launch is calculated for the high and low values of each parameter, the parameters are ordered (highest to lowest) based on the relative impact on the overall cost per launch, and the swings in cost per launch are plotted on the graph.

At the top of the graph are the parameters that have the greatest effect on the cost per launch, with positive effect in blue (for example, reducing cost) and negative effect in red. Parameters shown lower on the graph have a smaller effect on cost per launch.

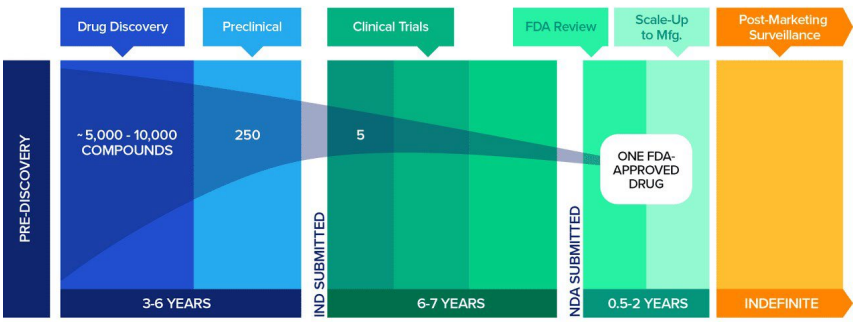
According to the tornado exhibit, more efficient preclinical research, could save hundreds of millions of dollars per drug launch. Additionally, ROIs could increase simply due to a more time efficient process.

So, for instance, if one assumes that AI plus extant databases are used to match the types of cells with the patients' cancers in a phase II trial, the chances of success might be much greater. This success rate is driven by improved comparability between the preclinical models and the clinical models (human tumors). On a per-launch basis, this could save about a billion dollars in R&D. According to the referenced analysis above, the impact of better quality, more efficient preclinical research, could save hundreds of millions of dollars per drug launch. Additionally, ROIs could increase simply due to a more time efficient process.

The probability assumption for moving through PII is 32.7% and through PIII to approval is 35.5%.

Failure rates in later stages of oncology research tend to be comparatively high. Phase I trials are often dose-escalation and safety analyses. Our research suggests that the accepted probability assumption for moving through PII is 32.7% and through PIII to approval is 35.5%. These PII and PIII to approval probability assumptions suggest that the preclinical models used to determine efficacy of a drug correlate extremely poorly with real human disease and/or heterogeneity of disease.

Exhibit 9: **Developing a new medicine takes 10-15 years**



Sources: Toptal Finance; UCSD Drug Development MOOC.

The Future of Drug Return on Investment

Average cost of drug discovery and development increased more than 82.49% from $1.188bn to $2.168bn (2010-2018).

In order to understand the importance and future of AI in pharmaceutical R&D, we must first frame the problems that are mounting for pharma's R&D programs and pipeline assets. According to Deloitte, the average cost of drug discovery and development increased more than 82.49% from $1.188bn to $2.168bn, per successfully launched drug, from 2010A through to 2018A.

Exhibit 10: **Average R&D costs, discovery to launch 2010-19A**

Sources: Deloitte Centre for Health Solutions. 2019. ""Ten Years on Measuring the Return from Pharmaceutical Innovation 2019 *The data in the exhibit uses original and extension cohorts

Probability of getting to drug launch after successful phase I trials is around 10%.

Deloitte estimates that 33% or $600-700m of total drug development and discovery costs are associated with the preclinical drug discovery phase per successful drug launch.

The average probability of moving from PII to PIII trials is <33% and from PIII trials to approval is little above 35%.

Average forecast peak drug sales are $407m YE18A vs. $816m Ye08A.

Market consensus probability of a drug getting to launch after successful phase I trials is around 10%. Therefore, after spending half a billion USD there remains a 90% chance of failure. According to Deloitte and others, average forecast peak drug sales had decreased to $407m YE2018A, which is less than half of the YE2008A average of $816m peak drug sales.

According to global consulting firm L.E.K.'s Launch Monitor, December 2020:

Average peak US drug sales over the last 15 years = $800m. 20% reach >$1bn and >50% reach <$250m peak sales.

•The average peak U.S. revenue of products launched over the past 15 years is about $800m, with only 1 in 5 reaching U.S. sales of $1bn, and over half failing to reach peak U.S. sales of $250m.

~50% of the peak sales values is reached by year three – a strong predictor of peak sales.

•On average, drugs products sales reach ~50% of peak sales by year three, and these early-year sales are strongly predictive of ultimate peak sales.

•About half of all drugs products launched over the past 15 years have underperformed pre-launch consensus forecasts by more than 20%.

Therapeutics for infectious disease, immunology and cardiovascular are the most prone to underperform market consensus estimates.

•Performance issues cut across all therapeutic areas, infectious disease, immunology and cardiovascular diseases, are the therapeutic areas where the most products underperform market consensus expectations.

Marketing power counts – on average large pharma peak drug sales values are 50% vs. smaller companies.

•On average, large pharma companies' product launches have average peak revenues 50% higher than those of smaller players.

•Small companies are also more likely to underperform expectations than larger companies, and this difference in revenue performance is magnified in diseases driven by primary care channels.

Stakeholder engagement and better forecasting are key to meeting expectations.

•Early launch planning, **effective stakeholder engagement**, better forecasting and highly disciplined execution can help address barriers and better manage market consensus expectations.

Late-stage pipeline candidate IRR has declined – from 10.1% in 2010 to 1.8% by 2018.

There has been a steady decrease in the IRR for late-stage pipeline candidates from 10.1% in 2010 to 1.8% in 2018, according to management consultancy, McKinsey.

Productivity of new launches peaked in 1997 at 3.1%.

Exhibit 11 below indicates the seven-year average revenue from new molecular entity (NME) launches divided by the portion of R&D spend over the preceding seven years. According to McKinsey, productivity of new launches peaked in 1997 at 3.1% and has been on a down trend since. A variety of sources and data support the McKinsey conclusions that ROIs are waning for pharma R&D, on average.

Reductions in both cost and time can reverse the decline in ROI vintage index.

ROI vintage index - Average productivity for drug R&D appeared to be in decline between 1998 and 2011, according to Mckinsey's ROI vintage index.

Can AI save enough cost and time?

What can improve the McKinsey metric is increased R&D efficiency, cost and time. We suggest that the best current hope for achieving better ROI in drug development via greater R&D efficiencies is AI driven drug discovery.

Uptick in the productivity trend (pre-AI) is likely to be misleading for investors.

During the period from 2012 to 2016, where the ROI vintage index relies more heavily on consensus market forecasts, there appears to be some indication that productivity could be improving. Research from L.E.K consulting (see above) causes us to infer that the uptick in the productivity trend is likely to be misleading for investors, because it is based more heavily on consensus market forecasts, which tend to have a 20% revenue overestimation bias.

In the exhibit below we detail the Mckinsey analysis for the ROI[1] vintage index from 1993 to 2016 and extend it using logarithmic and straight line trend forecasts to 2022E – The ROI vintage index formula is 7 years of revenues from NME launches[2,4] in a given year / by the portion of R&D spend over the preceding 7 years corresponding to a given vintage[3,4].

Annotations

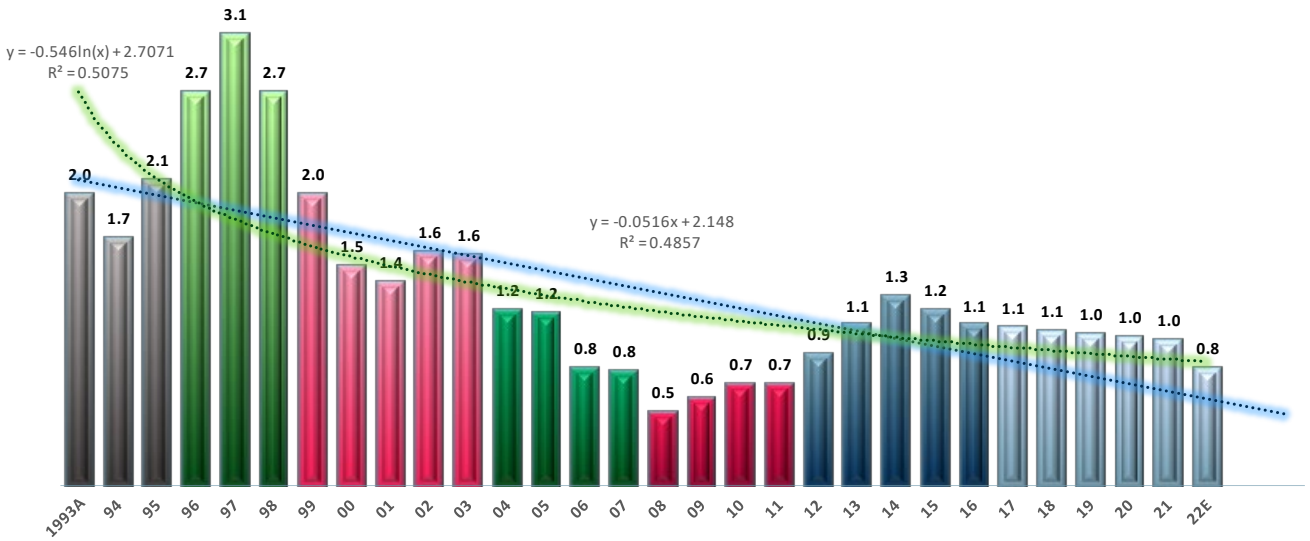
[1]Return on investment

[2]NME = new molecular entity; NME-grade products, excluding generics, biosimilars and new-drug-application products (i.e., new derivatives, reformulations, etc.); launch year based on global market entry and first reported/expected revenues; 3-year rolling average.

[3]Assigned based on average R&D progression and the proportion of spend attributed to different R&D stages.

[4]Inflation-adjusted to 2017 USD; sales beyond 2016 are based on analysts' forecasts.

Exhibit 11: **Productivity improving but uncertainty for future**



Sources: ACF Equity Research Graphics & Estimates Digital in R&D: The $100 billion opportunity. Evaluate Pharma May 2017; Pharmaceutical Research and Manufacturers of America 2016; McKinsey & Company analysis, 2017.

AI Iterative Active Learning in Drug Discovery

ROI in drug discovery and development is influenced by the failure of drug candidates to live up to pharmacological and market expectations.

A significant factor affecting ROI in drug discovery and development is the failure of drug candidates to live up to pharmacological and market expectations as consecutive tests are carried out i.e., as data is amassed.

Drug testing is carried out on humans because animal models are not necessarily true parallels.

Even after a drug is discovered, tested in preclinical (nonhuman) trials, and moved on to a clinical trial where it is tested on humans, it doesn't necessarily mean the drug will work as expected on humans - efficacy or safety or both, often do not live up to expectations.

Drug testing on humans remains a potentially very risky phase in a drug's development, potentially company destroying and lethal to human trialists.

A notable extreme and distressing example of continued testing and reducing expectations is the drug candidate TGN1412 designed to treat leukemia. TGN1412 was tested in monkeys where it was well tolerated. Unfortunately, when just 1/500th of a dose was given to six healthy young men in the first phase of clinical trials in 2006, they immediately developed fever, vomiting, and diarrhea. Within hours, the test group was referred to an intensive care unit with multiple organ failures. Drug R&D is a long and expensive process filled with unexpected failures, ripe for the help of AI.

According to a 2011 paper by QM/Predictive Oncology's Dr. Murphy:

Dimensionality of the experimental space is at the core of the risks associated with drug testing.

"At a fundamental level, the central problem of screening for potential drugs is the dimensionality of the experimental space within which screening takes place."

The number of experiments required to directly screen for compounds that affect one target while not affecting others can quickly become intractable.

Challenges include incomplete signaling pathway knowledge and the complexity of pathway interactions.

The only practical solution is to carry out a subset of the possible experiments. Current approaches in drug development require scientists to choose a path through experimental space guided by existing knowledge (e.g., signaling pathways), investigator insight and intuition. This process is often hindered by incomplete or incorrect pathway information and the difficulty of making predictions about complex pathway interactions.

A statistical model building approach using machine-learning (AI) looks a better solution.

An alternative described here (2011 paper by QM/Predictive Oncology's Dr. Murphy) uses active machine-learning (AI) methods to build statistical models of the entire space and iteratively choose experiments that are expected to improve the model.

Machine-learning (AI) allows for a purely empirical approach.

The major strength of this [iterative] approach is that experiment choice is guided on a purely empirical basis and in full consideration of the potential complexity of the system. Active learning is well established in some domains, and it has been applied in a few cases to biological problems.

Active learning iteration using AI is an interpretation of the empirical approach.

A subsequent CMU paper, published in 2014, Kangas, Naik, and Murphy, describes the 'active learning' AI iterative approach:

"Active learning consists of three phases performed in a loop…" (as illustrated in the exhibit 12 below).

A campaign of experiments can be initialized either using prior results from literature or databases or by randomly selecting a batch of experiments from an experimental space. The CMU paper outlines the experimental campaign design process as follows:

1) A model is generated to represent the currently available data.

(2) From that model, experiments are selected for execution that are expected to improve the model.

(3) The set of experiments is executed, and the resulting data are combined with previously collected experimental data.

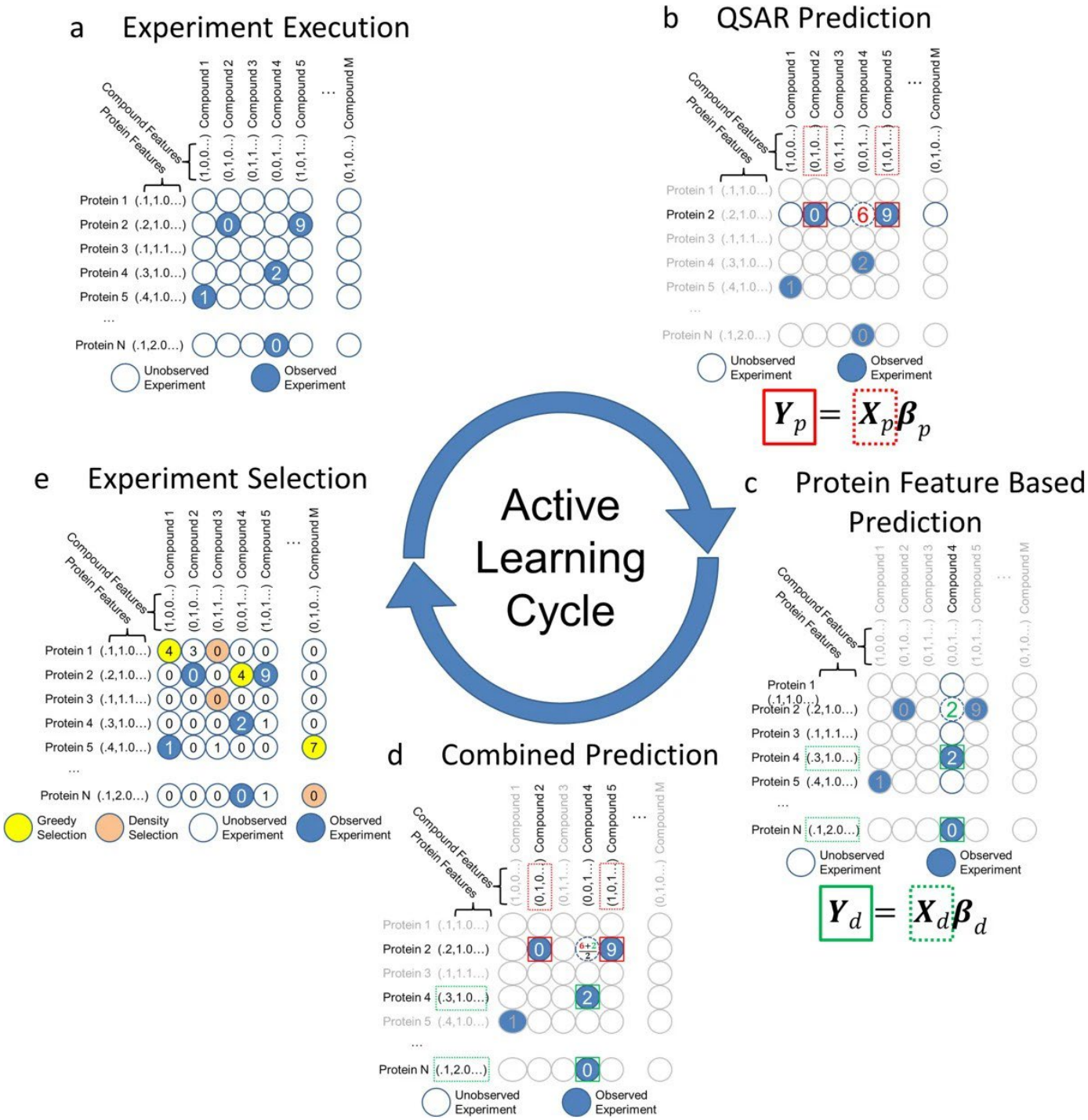
The loop then recommences from step 1 until either a desired accuracy of predictions is achieved, or a specified budget has been exhausted.

The CMU paper points out that there have been limited previous applications of AI driven active learning to the drug discovery process. In these efforts, compound activity was considered to be binary (active or inactive), and effort was focused on only a single target.

In exhibit 12 below we show an active learning pipeline for an experimental space with N proteins and M compounds. We provide a detailed description of the active learning iterative cycle approach.

In summary, researchers created two separate models, one using compound features only (CFO), and one using protein features only (PFO), then averaged the two for selection of experiments. The experiments are moved forward by targeting information where experimental success is most likely or where there is the greatest information void.

Exhibit 12: **AI active learning cycle in drug discovery**



Sources: Kangas, J. D. (2014). Efficient discovery of responses of proteins to compounds using active learning. BMC Bioinformatics, 15(1), 143.

(a) Experiment Execution - A round of active learning begins with the data for all of the experiments that have been observed so far.

(b) QSAR Prediction - A separate model is constructed for each protein using the compound features to make predictions for the effect of each compound on the activity of that protein. This is illustrated for Protein 2, for which regression using the observed experiments for Compounds 2 and 5 predicts that Compound 4 would show an activity of 6. This model is referred to as Compound Features Only (CFO), which is analogous to QSAR, which checks for the presence or absence of specific structural elements in the drug compound, to predict interactions between endogenous proteins and drugs.

Quantitative structure-activity relationship (QSAR) – a computational method for examining relationships between chemical compound structure properties and biological actions.

(c) Protein Feature Based Prediction - A separate model (molecular docking model) is constructed for each compound using the protein features to make predictions for the effect of that compound on the activity of each protein (which requires information on both the protein target and the drug). This is illustrated for Compound 4 for which regression using the observed experiments for Proteins 4 and N predicts that Protein 2 would show an activity of 2. This model is referred to as Protein Features Only (PFO).

Molecular docking is the study of how two or more molecular structures fit together (i.e., drug and enzyme protein).

(d) Combined Prediction - For the CCT approach, if predictions from both methods are available, they are averaged. In the early rounds when no experiments may have been observed for a given protein or compound, predictions from both models may not be possible.

(e) Experiment Selection - The complete set of observations and predictions is shown, and experiments that would be chosen for the next round of acquisition by different methods are shown (greedy selection would pick the experiments with the highest predicted values, while density selection would pick experiments for compounds and proteins that are most different from those previously selected). The results for the chosen experiments are added to those experiments observed so far to begin the next round of active learning.

Inference - AI driven active learning generates the greatest amount of useful data while minimizing the amount of experimentation (time and money) required to obtain that information.

Active learning driven by AI permits simultaneously modelling of the effects of multiple compounds on multiple targets.

As described above the innovation achieved via active learning for drug discovery and development when compared with previous approaches is the emphasis on machine learning to simultaneously model the effects of many compounds on many targets.

Could active-learning lead to a 24x increase in the rate of compound-target pairs discovery?

Predictive Oncology (NasdaqCM : POAI) claims that that active compound-target pairs could be discovered as much as 24x faster using active learning than by random selection of experiments. With appropriate algorithms, the process can be applied to large experimental spaces.

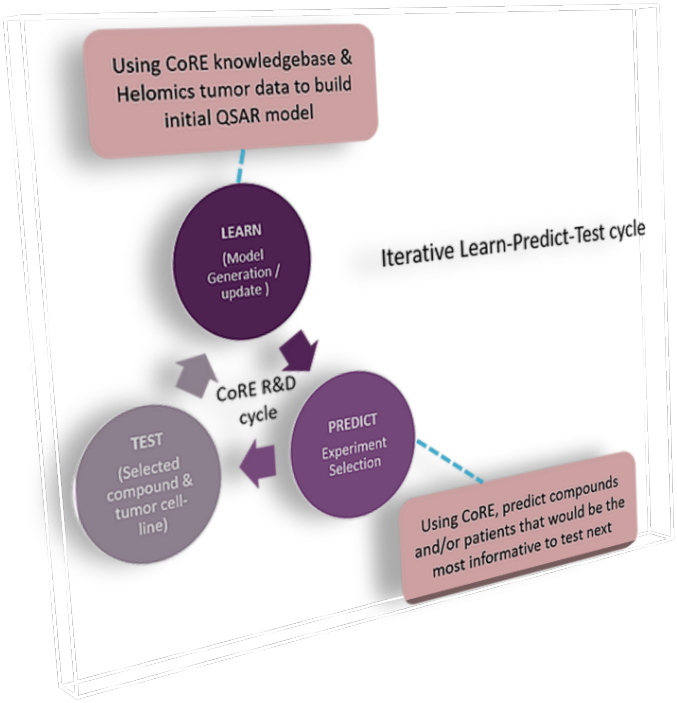
Based on data and observations from AI drug discovery companies we infer that AI driven active learning cycles can:

1) successfully reduce compound synthesis required to discover promising drug leads;

2) reduce experimentation cost to develop accurate predictive models compared to industry-standard approaches;

3) reduce experimentation cycles by using historical experimental results.

Exhibit 13: **POAI's approach to drug discovery – active learning**



Sources: ACF Equity Research Graphics; Predicative Oncology (Helomics).

AI and 3-D cell culture media

We infer investment opportunities in media design and composition using AI.

AI could be leveraged by pharma companies in several areas of drug discovery and development, helping them reduce their development cost and time while improving accuracy and predictability.

AI could be deployed to improve human tumor growth. Human tumors are challenging to grow in a cell culture flask, and even if they are grown, a typical immortal cell line culture does not imitate the original tumor.

Over time, cancer cell lines have adapted to life in plastic petri dishes, altering their genetic make-up and behaviour.

In vitro cancer cell lines have lived over many generations in an environment that differs radically from their native one. Over time, the cells have adapted to life in plastic petri dishes, altering their genetic make-up and behaviour. The drive now is to create models in which the tumors grow in an environment that, although not human, better mimics their native environment.

Therefore, we infer that progress opportunities are available in media design and composition.

Technologies that help preserve the patient's derived cancer tissue's biological signature allow researchers to study cancer in preclinical models that should mimic real tumors much more closely.

AI driven drug development could create better models for culturing and screening patient-derived cell lines (PDCL).

Immortal cell lines that better mimic real tumors should hopefully translate into scientific findings that translate into better clinical outcomes. The objective in AI driven drug development is to create better models for culturing and screening patient-derived cell lines (PDCL), which will help with scientific advances at the preclinical level as well as raising clinical success rates.

UK cancer researchers using improved media mixes (now available globally) independently isolated, identified a unique set of 71 new ovarian cancer cells from patients. Adding these 71 cell lines to an existing 25 cell lines created a powerful new resource for researchers. Of the original 25 cell lines, 11 were used to represent around 95% of ovarian tumors.

Media is available that could retain >95% of cell DNA, RNA and proteomic signatures over far greater expansion cycles.

The media used to grow the new unique ovarian cancer cell lines, licensed by Predictive Oncology and its partner, GLG Pharma, is able to retain 95%+ of the cell's DNA and RNA as well as crucial proteomic signatures, even after 70 expansion cycles, as opposed to the standard 20 cycles. These cell lines generally differ significantly from standard cell lines.

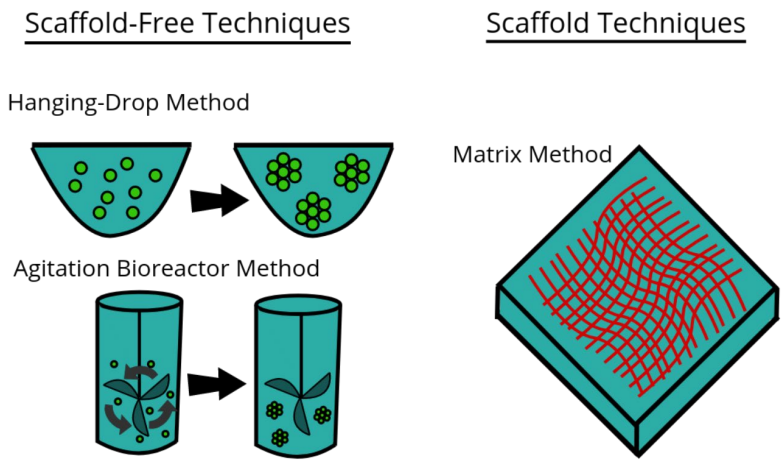
New media approaches (AI driven) may create a revolution in cancer cell line usefulness.

Ovarian cancer cell lines are often collected from ascites fluids, which are notoriously difficult to work with and when using standard (historical) media mixes, are often prone to failure. In addition, these ascites fluid derived cell lines, are often not representative of the patient's unique ovarian cancer. New media approaches may create a revolution in ovarian cancer cell line usefulness.

3D cancer cell culture media market estimated value2027E $3.2bn, CAGR >11%.

The 3D cancer cell culture media market is expected to grow at a CAGR of approximately 11.3%, to $3.2bn worldwide in sales in 2027, according to Allied Market Research. Grandview research estimated that the global 3D cell culture market had a value of $1.5bn in 2020 and expects a 10.7% CAGR to YE28E.

Exhibit 14: **Examples of 3D media environments**

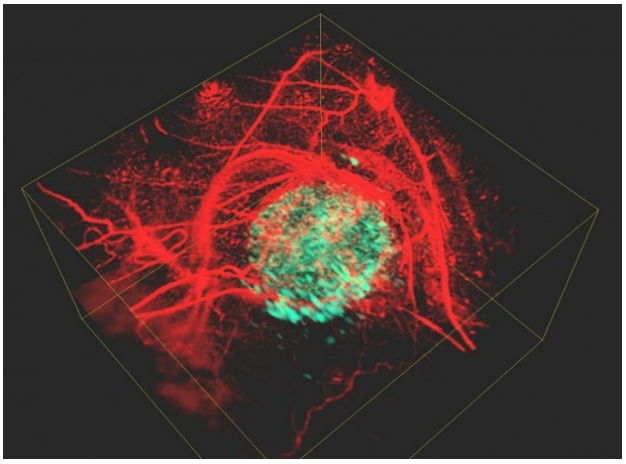


Sources: Wikipedia commons.

90% of breast and ovarian cancer research is carried out using 10 standard cancer cell lines (SCCL). These SCCLs have often lost relevant biomarkers.

90% of published cell line research in breast and ovarian cancer is carried out with ten standard cancer cell lines (SCCL). Many of these SCCL have lost the relevant lineage biomarkers for the breast and ovarian tissues in question. The molecular profile of the SCCL xenograft frequently differs from the primary tumor.

Exhibit 15: **3D images of cancer cells in vivo**



Sources: Jan Laufer; Healthcare-in-europe.com.

AI and Tumor Heterogeneity

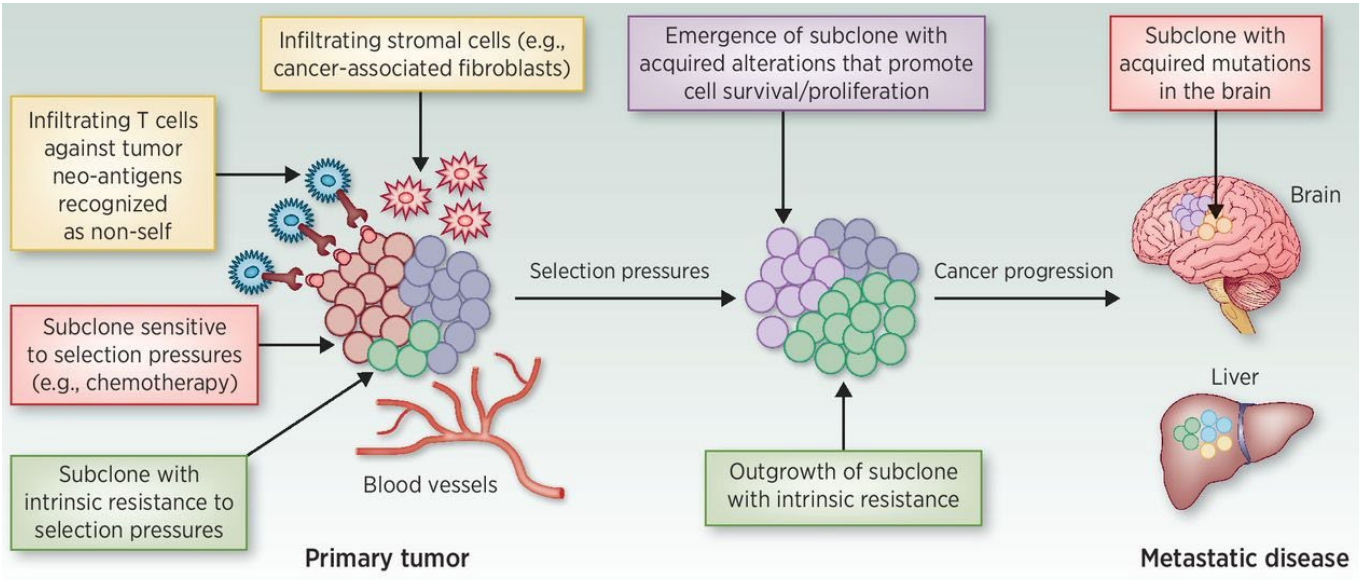
Cell lines diversification is critical because of tumor heterogeneity.

Tumor heterogeneity has been identified as a major clinical cancer treatment issue since around 1980.

Clearing one subset of tumor cells provides only temporary patient benefits.

A key problem in cancer research and clinical practice is tumor heterogeneity. A primary problem in cancer treatment is the inherent heterogeneity of cancers, which translates to an inability to destroy all the different types of cancer cells within a tumor and its metastases. Tumor heterogeneity was first identified and proposed as a major issue in clinical practice in ~1980. The exact mechanisms by which tumors become heterogenous is still consistently debated. Regardless, the translation into the clinic is fairly simple: clearing only one subset of tumor cells only temporarily provides benefit for the patient.

Exhibit 16: Tumor heterogeneity emergence



Sources: CCR Reviews; American Association for Cancer Research © 2005; Jamal-Hanjani, M et al. (2015). Translational Implications of Tumor Heterogeneity. Clinical Cancer Research, 21(6), 1258-1266.

Chemotherapy destroys rapidly dividing cells – this is often only a creates a temporary stasis in disease development.

Tumor heterogeneity is the bane of chemotherapy. Many chemotherapies are designed to destroy rapidly dividing cells (which quickly grow into a tumor), and therefore stop the cancer in its track. Often, destroying rapidly dividing cells only provides temporary respite from the development of the disease.

Chemotherapy effects are often temporary because it creates selection pressures, killing only the most aggressive cells.

The effects are temporary because chemotherapy creates a selection pressure by killing only the most aggressive cells. It also acts to make way for cells that have survived the chemotherapy and so are resistant to that chemotherapy, or it provides residual cells, previously outcompeted (for nutrients) by the more aggressive cells, with the ability to grow.

A clinical regimen that could identify all the different cell variants in a tumor and administer a drug, or multiple drugs, which renders all the cells susceptible could improve clinical outcomes, and in the perfect outcome, eradicate the cancer permanently.

From the above observations, we infer that it is necessary to use:

a) PDCLs - preclinical research that can use PDCLs that are heterogeneous and replicate the patient tumor and;

b) AI - preclinical research that uses AI to identify drugs or drug cocktails that address various kinds of tumor heterogeneity.

AI combined with PDCLs could transforms the number of agents and biomarkers available to treat a cancer type.

Combined PDCLs and AI use in preclinical research may deliver more success when translated into clinical studies, compared to single agents used in conjunction with only a handful of biomarkers.

PI trial successes in oncology are higher than PII success rates (below 30%).

Other analyses suggest PI trial successes in oncology are higher than PII success rates (below 30%), further indicating poor translation of preclinical efficacy to clinical efficacy. Thus, companies in the AI drug development space with assets (AI, database(s) and PDCLs) could be extremely useful in increasing the chances of success in mid-and-later stages of research (PII and PIII).

Only one out of between 5,000-10,000 drugs are able to make it to FDA approval.

Overall, only one out of between 5,000-10,000 drugs are able to make it to FDA approval—accurate preclinical research could drastically increase chances of success and investor ROI.

A Peer Group - AI Drug Development

Exhibit 17: Peer group selection of AI drug development companies

TTM Metrics / Company Name	Market	Tkr	Price	MCAP US$(m)	EV $(m)	52 Wk Chg %	P/ sales	EPS Dil	PER	Staff	Rev / Per Head
Exscientia	XNAS	EXAI	8.91	1,097	502.03	-66.94%	26.48x	-1.33	N/M	287	125,087
BioXcel Therapeutics	XNAS	BTAI	13.27	352	181.12	-56.02%	N/M	-4.11	N/M	89	N/M
e-Therapeutics	XLON	ETX	21.94	113	83.27	-33.25%	196.82x	-1.94	N/M	32	17,011
Lantern Pharma	XNAS	LTRN	5.01	54	-7.76	-55.82%	N/M	-1.26	N/M	14	N/M
Predictive Oncology	XNAS	POAI	0.39	28	0.50	-62.50%	16.22x	-0.31	N/M	30	49,333
Average					152	-54.9%	79.84x	N/M	N/A	90	63,810
Median					83	-56.0%	26.48x	N/M	N/A	32	49,333

Sources: ACF Equity Research Estimates

TTM Metrics / Company Name	Market	Tkr	MCAP US$(m)	Gross Debt / Assets	Gross Debt / to Equity	RoA	RoE	RoI	P/ sales	P/ book	β 5yr
Exscientia	XNAS	EXAI	1,097	0.29	1.25	-0.14	-0.18	N/M	26.48x	1.78x	N/M
BioXcel Therapeutics	XNAS	BTAI	352	0.01	37.94	-0.29	-0.57	N/M	N/M	2.14x	1.22x
e-Therapeutics	XLON	ETX	113	0.00	2.45	-0.27	-0.39	N/M	196.82x	3.54x	0.38x
Lantern Pharma	XNAS	LTRN	54	0.02	0.21	-0.14	-0.23	0.12	N/M	0.89x	1.01x
Predictive Oncology	XNAS	POAI	28	0.00	1.60	-0.18	-0.62	N/M	16.22x	0.84x	1.32x
Average				6.53%	8.69%	-72.47%	12.44%	0.12%	79.84x	1.84x	0.98x
Median				0.72%	1.60%	-38.39%	12.44%	0.12%	26.48x	1.78x	1.12x

Sources: ACF Equity Research Estimates

TTM Metrics / Company Name	Market	Tkr	MCAP US$(m)	EV $(m)	MCAP / REVS	MCAP / EBITDA	MCAP / NI	MCAP / FCF	Rev / Per Head	EBITDA / Per Head	FCF / Per Head
Exscientia	XNAS	EXAI	1,097	502.03	30.56x	N/M	N/M	N/M	125,087	-342,683	-108,815
BioXcel Therapeutics	XNAS	BTAI	352	181.12	N/M	N/M	N/M	N/M	N/M	-1,352,809	-684,944
e-Therapeutics	XLON	ETX	113	83.27	207.41x	N/M	N/M	N/M	17,011	-339,864	-208,626
Lantern Pharma	XNAS	LTRN	54	-7.76	N/M	N/M	N/M	N/M	N/M	-1,150,714	-528,571
Predictive Oncology	XNAS	POAI	28	0.50	19.07x	N/M	N/M	N/M	49,333	-415,333	-275,667
Average					85.68x	N/A	N/A	N/A	63,810	N/M	N/M
Median					30.56x	N/A	N/A	N/A	49,333	N/M	N/M

Sources: ACF Equity Research Estimates

Glossary

AI	AI is the abbreviation for so-called Artificial Intelligence. Previously used to describe machines that mimic "human" cognitive skills associated with human intelligence, e.g. learning and problem-solving. Now more commonly thought of in terms of machine-based rationality and acting rationally, removing a former limit to accepted possibilities for AI. Common examples of current AI applications and capabilities include advanced web search engines such as Google, recommendation systems e.g. Netflix, understanding human speech e.g. Alexa. See also machine learning.
Apoptosis	Apoptosis is also referred to as programmed cell death. It is the normal planned and controlled death of cells as a function of the growth and development of an organism.
CAGR	CAGR is the compound average growth rate over a certain time period calculated by taking the value at the beginning or the series and the end of the series and using the number of periods 0 (starting value)-X (finishing value) to establish an average growth rate per period.
CoRE™ active learning	CoRE (Computational Research Engine) active learning is a comprehensive in silico platform that iteratively optimizes predictive models using guided selection of experiments developed at the Ray and Stephanie Lane Center for Computational Biology at Carnegie Mellon University (source https://helomics.com/ai-pharma/core-active-learning/)
EBITDA	Earnings before interest, depreciation and amortisation – the presentation of EBITDA by companies is not a requirement of UK GAAP or IFRS accounting standards. However, in certain cases it can act as a close proxy to free cash flow.
EBT	Earnings before tax. Also expressed as PBT – profit before tax.
FCF	Free Cash Flow generated in ACF's models after all obligatory cash costs have been satisfied such as Interest payable (Ip), cash taxes and maintenance capex (as opposed to investment capex). FCF represents the cash remaining for theoretical distribution or investment after all obligatory cash-based costs including net interest payable have been deducted.

FDA approval	FDA approval means a company has demonstrated that its drug or biological product is safe and effective for the intended use, and that it can manufacture the product to federal quality standards (source: www.fda.gov).
Hormetic dose response	Hormetic dose response describers the dose response phenomenon to xenobiotics (toxins) or other stressors in which an individual can derive positive benefits at certain concentrations of the toxin. Hormesis is characterized by a low-dose stimulation, with zero dose and a high-dose inhibition describing a J-shaped or an inverted U-shaped dose response. In these graphs the arms of the "U" are inhibitory or toxic concentrations and the curve regions show stimulation of a beneficial response.
JV	Joint Venture – generally a legal structure between two corporate entities involving participation in equity capital in the JV vehicle. JV can also refer to more informal arrangements.
Machine learning	Machine learning (ML) is a field of research that resolves problems related to the understanding and building methods that 'learn'. It can also be described as methods that leverage data to improve performance on some set of tasks. It is generally considered a skill set within the field of artificial intelligence (AI). How it works - Machine learning algorithms build a model based on sample data (the training data), then make predictions or decisions without explicit instructions to do so. See also AI (artificial intelligence).
NoSh	Number of Shares in issue (NoSh).
NPV	Net Present Value (NPV) refers to the current value of future cash flows generated by the project
PDCL	PDCL - Patient-derived cell lines.

PeDAL™	Patient-centric drug discovery using active learning (PeDAL) is a unique technology that combines a proprietary, clinically validated patient tumor cell line assay, a vast knowledgebase of proprietary and public data, and Active Learning – the Active Learning allows for the efficient exploration of compound drug responses against a large diverse patient "space". PeDAL offers researchers the opportunity to bring patient diversity efficiently and cost-effectively into drug discovery much earlier. (Source https://helomics.com/ai-pharma/pedal/).
ROI	ROI or Return on Investment is a financial metric used to measure or compare the relative efficiency of one investment versus other potential investment opportunities. ROI's central concept is to divide an investment's net profit (loss) by its initial cost. It is generally only applicable as a comparative returns tool when comparing two or more projects that are similar or share key characteristics.
SCCL	SCCL - Standard cancer cell lines.
WACC	Refers to the weighted average cost of capital for the firm.
QSAR	QSAR - Quantitative structure-activity relationship is a computational modeling method for revealing relationships between structural properties of chemical compounds and biological activities
3D cell culture	A 3D cell culture is an in-vitro environment in which biological cells grow / interact in all three dimensions. Unlike 2D environments (e.g. a Petri dish), a 3D cell culture allows cells in vitro to grow in all directions, and is designed to better replicate how cells grow in vivo. 3D cell cultures are constructed using scaffold techniques with hydrogels and other materials.

INDEPENDENCE & DISTRIBUTION

ACF Equity Research Ltd is a provider of issuer-pays research with a clearly defined independent ethic. ACF produces accurate, clear, focused research aimed at a professional investment audience. ACF has excellent distribution capabilities and always aims to provide access without restriction to the widest professional audience. ACF offers a range of additional services to support its clients.

DISCLAIMER

This communication is for informational purposes only. It is not intended as an offer or solicitation for the purchase or sale of any financial instrument or as an official confirmation of any transaction. [The opinions expressed in this report herein do not take into account individual investor circumstances, objectives, or needs and are not intended as recommendations of particular securities, financial instruments or strategies to particular clients. The recipient of this report must make its own independent decisions regarding any securities or financial instruments mentioned herein.] ACF Equity Research Ltd has based this document on information obtained from sources it believes to be reliable, but which it has not independently verified. Neither ACF Equity Research Ltd. nor any of its directors, officers, employees or agents shall have any liability, however arising, for any error, inaccuracy or incompleteness of fact or opinion in this research report or lack of care in this research report's preparation or publication, or any losses or damages which may arise from the use of this research report. All market prices, data and other information are not warranted as to completeness or accuracy and are subject to change without notice. [Past performance is not necessarily a guide to future performance and no representation or warranty, express or implied, is made by ACF Equity Research Ltd. with respect to future performance. Any recommendation or opinion contained in this research report may become outdated as a consequence of changes in the environment in which the issuer of the securities under analysis operates, in addition to changes in the estimates and forecasts, assumptions and valuation methodology used herein.]

IMPORTANT DISCLOSURES FOR U.S. INVESTORS

This research report was prepared by ACF Equity Research Ltd., a company authorized to engage in securities activities in the United Kingdom. ACF Equity Research Ltd. is not a registered broker-dealer in the United States and, therefore, is not subject to U.S. rules regarding the preparation of research reports and the independence of research analysts. This research report is provided for distribution to "major U.S. institutional investors" in reliance on the exemption from registration provided by Rule 15a-6 of the U.S. Securities Exchange Act of 1934, as amended. Any U.S. recipient of this research report wishing to effect any transaction to buy or sell securities or related financial instruments based on the information provided in this research report should do so only through, a registered broker-dealer in the United States. The analyst who prepared this research report is not registered or qualified with the Financial Industry Regulatory Authority ("FINRA") and may not be associated with a U.S. broker dealer and as such, would not be subject to applicable restrictions under FINRA Rules on communications with a subject company, public appearances and trading securities held by a research analyst account.

[Investing in any non-U.S. securities or related financial instruments (including ADRs) discussed in this research report may present certain risks. The securities of non-U.S. issuers may not be registered with, or be subject to the regulations of, the U.S. Securities and Exchange Commission. Information on such non-U.S. securities or related financial instruments may be limited. Foreign companies may not be subject to audit and reporting standards and regulatory requirements comparable to those in effect within the United States.]

LEGAL NOTICE

This report is for authorized use by the intended recipient(s) only. It may contain proprietary material, confidential information and/or be subject to legal privilege. No part of the content of this research report may be copied, forwarded or duplicated in any form or by any means without the prior consent of ACF Equity Research Ltd. and ACF Equity Research Ltd. accepts no liability whatsoever for the actions of third parties in this respect.

IMPORTANT ADDITIONAL DISCLOSURES FOR U.K. INVESTORS

The information in this report has been prepared by ACF Equity Research Ltd (ACF). The research is published for information purposes only. It is not intended as an offer or solicitation for the purchase or sale of any securities or any financial instruments.

ACF has based this document on information obtained from sources it believes to be reliable but which it has not independently verified. All market prices, data and other information are not warranted as to completeness or accuracy and are subject to change without notice. Any comments or statements made herein do not necessarily reflect those of ACF Equity Research Limited. The material should not be regarded by recipients as a substitute for the exercise of their own judgment. Past performance does not guarantee future performance.

The analyst(s) responsible for covering the securities in this report receive compensation based upon, among other factors, the overall profitability of ACF which may, from time to time, solicit business from any of the companies referred to in this report. The analyst(s) responsible for covering securities in this report may not hold a position in any or related securities in this report in ACF's sector universe on in any other sector in which ACF carries out research. The company does not hold any position in the securities mentioned in this report.

This research report and its contents are intended for professional investors and not for retail investors. A marketing communication under FCA Rules, this document has not been prepared in accordance with the legal requirements designed to promote the independence of investment research and is not subject to any prohibition on dealing ahead of the dissemination of the investment research. ACF Equity Research Limited is authorised and regulated by the Financial Conduct Authority. However the contents of this research report are produced as if ACF Equity Research Limited is unregulated and consequently this report does not contain investment recommendations or ratings.

ACF, its directors, employees and agents accept no liability whatsoever for any loss or damage of any kind arising out of the use of all or part of these materials. The information in this report is provided with the understanding that ACF is not acting in a fiduciary capacity. Certain laws and regulations impose liabilities which cannot be disclaimed. This disclaimer shall in no way constitute a waiver or limitation of any rights a person may have under such laws and/or regulations. Copyright © 2018 ACF Equity Research all rights reserved. Additional information is available upon request.

ACF Equity Research Limited, 125 Old Broad Street, London, EC2N 1AR, U.K.
Tel: +44 (020) 7558 8974
Website: www.acfequityresearch.com