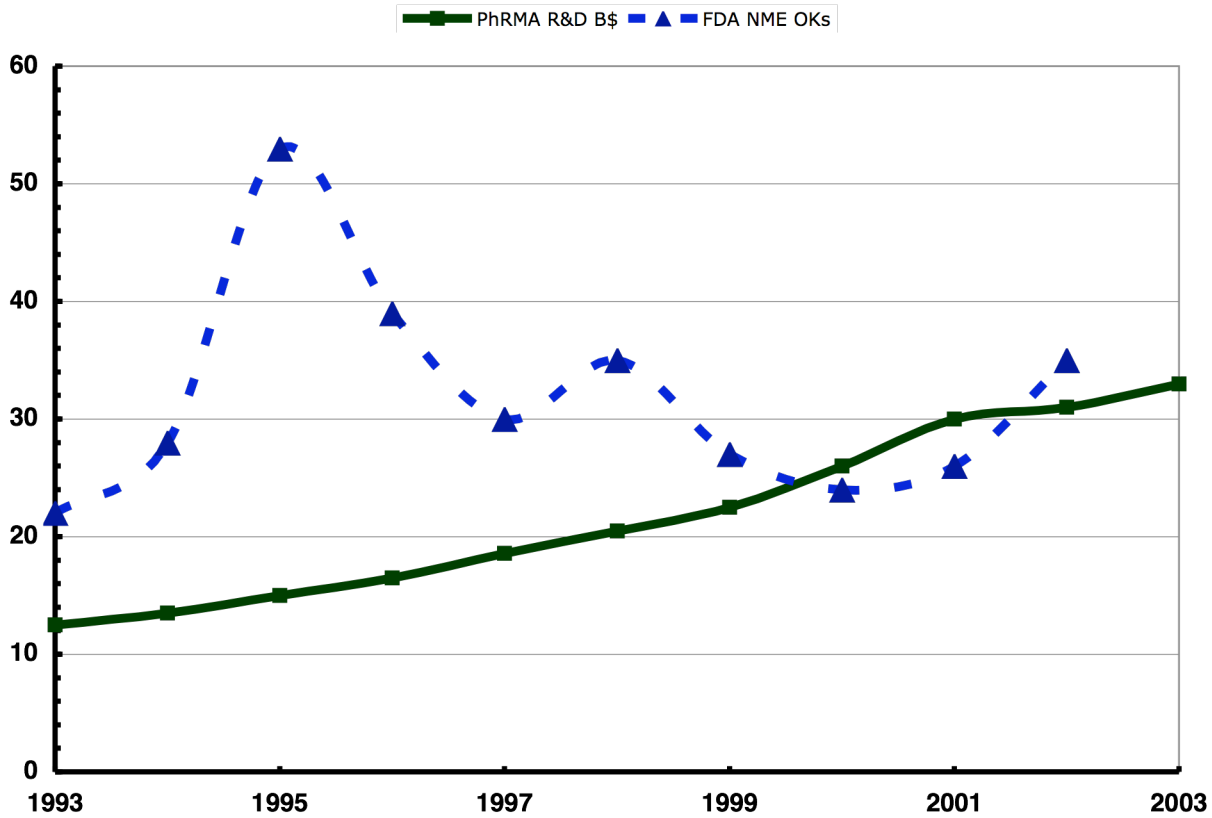


Translational and Critical Path Research
Project Excellence, Total Quality, Strategic Coherence, Tactical Brilliance

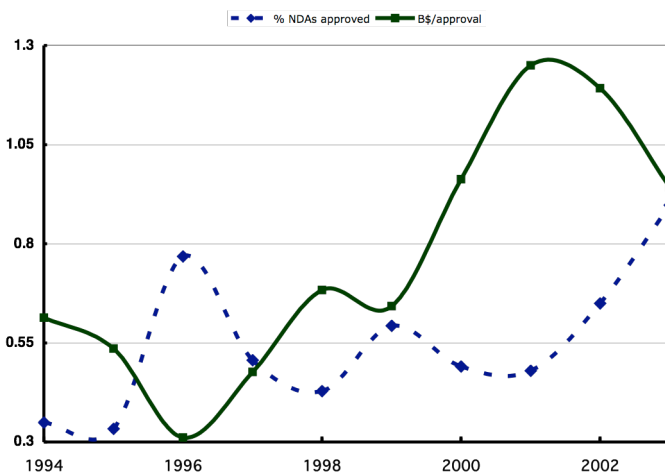
Translational research, from discovery to market, has become an urgent focus of U.S. biomedical research efforts. Elias Zerhouni, MD, Director of NIH, has set Translational Research as a priority and many institutes have formed offices and programs to facilitate the efficient progression from the laboratory to market¹. The FDA believes: *...the applied sciences needed for medical product development have not kept pace with the tremendous advances in the basic sciences...Not enough applied scientific work has been done to create new tools to get fundamentally better answers about how the safety and effectiveness of new products can be demonstrated, in faster time frames, with more certainty, and at lower costs. In many cases, developers have no choice but to use the tools and concepts of the last century to assess this century's candidates. As a result, the vast majority of investigational products that enter clinical trials fail...the path to market even for successful candidates is long, costly, and inefficient, due in large part to the current reliance on cumbersome assessment methods. A new product development toolkit—containing powerful new scientific and technical methods such as animal or computer-based predictive models, biomarkers for safety and effectiveness, and new clinical evaluation techniques—is urgently needed to improve predictability and efficiency along the critical path from laboratory concept to commercial product. We need superior product development science to address these challenges...The medical product development process is no longer able to keep pace with basic scientific innovation...A third type of scientific research is urgently needed, one that is complementary to basic and translational research, but focuses on providing new tools and concepts for the medical product development process...we will call this critical path research.*² Similar translational research initiatives have been started in the EU and UK.

History: In the 1970s the fully allocated cost of research associated with launch of a new chemical entity (NCE) was <\$100 million. In the 1980s as costs for big pharma escalated, investors saw in the miraculous new molecular/cell biology an opportunity to exploit human peptides and monoclonal antibodies (MoABs) for faster, lower cost, lower risk products with market attractiveness and premium pricing. A bonanza—it lasted less than a decade. A few got rich. Many failed. Reality set in. In the 1990s biotech products merged with traditional NCEs in regard to costs, timing, and value and their probabilities of success converged. In the 2000s all have run into serious problems³. Exploring new complex, multifactorial, polygenic disorders with single unknown pharmacologies has led to increasing failures. Again and again an unstratified pivotal study fails, *post hoc* analysis suggests some subgroup might be successful, and confirmatory studies prolong the agony without a good result. One successful pivotal study often is not replicated. The FDA, smug in its misinterpretation of congress's meaning of *two studies* (they meant more than one single investigator) now reaffirms the need for extensive predefined triple-blinded replicated pivotal studies, leading to many biotechs becoming *after the hyphen* or just obliterated.

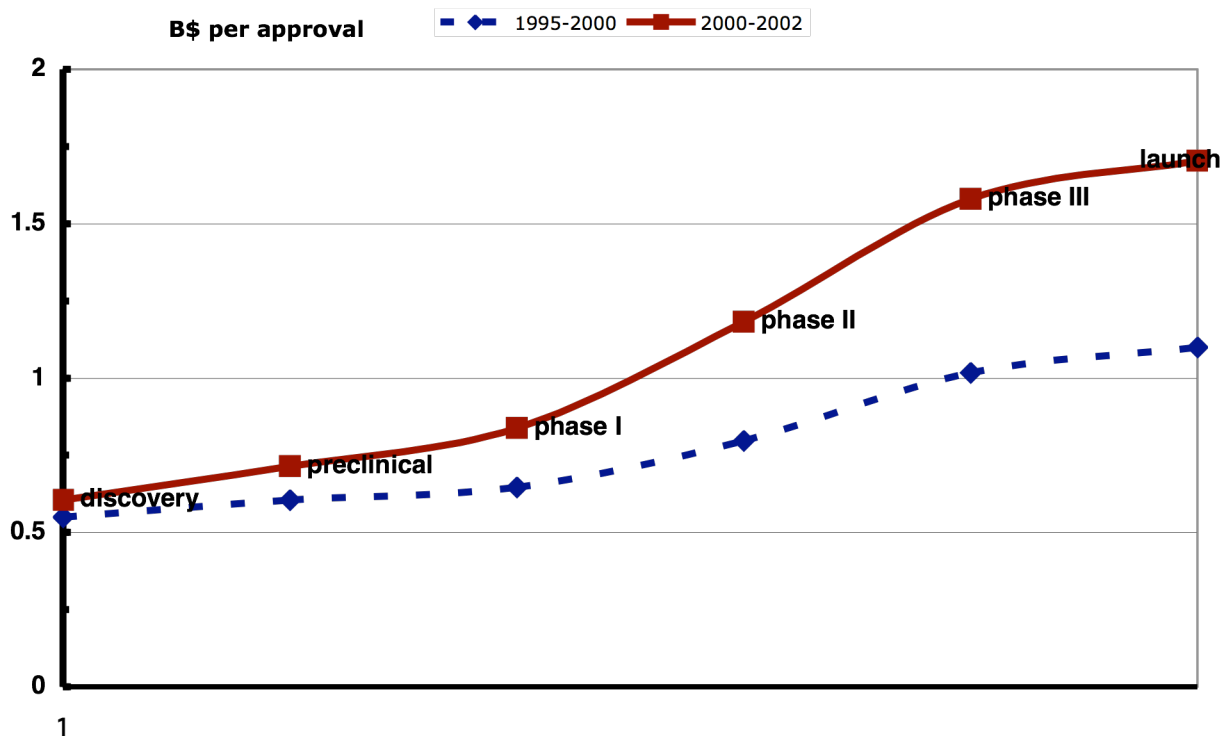
Pharmaceutical research investments are growing. The chart below⁴ tracks the research investments only by members of the U.S. trade association, PhRMA, and FDA approvals of NCEs including vaccines and biologics. Most of the significant NCEs approved world wide are approved by FDA but PhRMA spending is probably only 2/3 of world research investments.



The figure below² shows the ratio of approvals each year to the total new applications for NCEs (including biologics and vaccines) in that same year. It also shows one estimate of the investments per approval using only the U.S. data from the first chart.



Another estimate of the investment per approval is from the Bain drug economics model for 2003 as shown below⁵. Note that the investment in discovery has changed little (10%), post-launch research increased 50%, but critical path (team formation to launch) increased 110%. Despite increased investment in this phase, the duration has not shortened, approval percentage has worsened, and costs have escalated⁶.



Current status: About one percent! of excellent discovery initiatives will reach market. Discovery is often prolonged by endless animal models replicated excessively for publication but not really predicting human efficacy or toxicity with the new disease targets and yet undefined biology. When a successful SAR yields one or several molecules to develop, there is a 5-10% probability of a launch. When the first human dose is achieved, the former 22% probability of a FDA approval is now about 10-15%. After proof of principle, when we used to expect at least an 85% probability of success, now half of projects fail during or after expensive pivotal trials⁷. Regulatory review is being accomplished rapidly and probably there is little more savings in time to be gained there. Regulatory documents have been simplified with a unified worldwide technical document and electronic submissions. But overall timelines have not improved and remain about a decade from team to launch. We are living on cheaper research done in years past, but facing price controls, new safety concerns, and substitution not just from generics and therapeutic equivalents but from cheaper markets.

Biotechnology differs: Big pharma continues to apply a diversified *shots on goal* approach across many therapeutic and technology platforms⁸. If they spend \$2 billion/year on research they expect at least one NCE launch and with global marketing/sales muscle they hope the scales will balance after a decade of sales. Biotechnology cannot pursue a multiple platform strategy as they can't raise the capital. If ten biotechnology firms can raise

\$200 million each, odds are that one will survive the first NCE launch but that is no guarantee of being immortal. The last big pharma to spring from discovery, Syntex, with steroids from Mexican yams, was gobbled up by a T rex to emerge only as droppings along the trail.

Picking winners: If biotechnology must focus on one or a very few platforms of diseases and pharmacology to have resources to be able to bring at least one success to fruition, it seems obvious they should pick the winners. Unfortunately, that is difficult. Tracking big pharma over four decades shows that one or another is dominant for a decade, but none have sustained success despite only pursuing exciting prospects. The dominant firm in the early years was Parke–Davis—where are they now? The strategy must be to avoid losers, to identify losers as quickly as possible (**fail fast**), and to be as efficient as possible. Although biotechnology science is very exciting, losers often result from competition from better molecules further along in development or with more effective marketing muscle. Losers also result from chinks in the intellectual property armor. Failing fast may be more difficult emotionally if you only have one platform, but it is still essential economically to avoid wasting resources and exposing patients unnecessarily to ineffective or unsafe agents. Achieving excellence in critical path research is essential and a recent study showed that firms that have achieved excellence have grown at 30.4% per year compared with 10.9% for those performing more sluggishly⁹.

Project Excellence: If we define a project as the critical path research between selection of a molecule or two from an exciting structure activity relationship series (SAR) to launch of a product on global markets, what are the principles of achieving excellence in this endeavor? The principles seem obvious and important and why they have been neglected is puzzling. First, select the right team to lead the effort. Second, plan effectively and minutely. Third, don't pursue unnecessary or low value activities. Fourth, align this project with the enterprise strategies.

Project Team: The team will drive the project and must collectively have profound knowledge and expertise of many *disciplines*, be effective in liaising with colleagues and collaborators, and be driven to succeed¹⁰. Obviously they must be bright and knowledgeable with a history of success, but they must also have personalities that allow effective teamwork. They should be selected, trained, mentored, and rewarded with recognition that their success may be the only way for the enterprise to succeed. The team often has seven key members, which is the number for most effective communication. A chemist must be able to represent analytical, bulk synthesis, formulation, and fill–finish worldwide. A biologist must be able to represent pharmacology and toxicology of all types. A clinician

must be able to represent clinical trials, investigators, safety problems, and ethical concerns worldwide. A regulatory affairs specialist must be able to interface with and predict regulatory behaviors worldwide. A kineticist must be an expert with mathematical models and be able to master kinetics in animals, toxicology, and diverse types of patients and master complex *in silico* models. A statistician/mathematician must be expert at trial design and analysis and in interfacing with statistical concerns of regulators and prescribers. Discovery, marketing, finance, and legal also need to be represented. The project team leader must span all of science, marketing, business, politics, ethics, and other domains as well as being an outstanding mentor, leader, motivator, and achiever. Many feel that collocation is important and a large common office space may facilitate teamwork and communication. Certainly the team must have cutting edge informatics technology and strive to be paperless.

Profound knowledge of the target: The first task of the team is to acquire profound knowledge of the target: the disease, alternative therapies, competitive products, intellectual property, and significant global markets. Is the disease defined the same in Japan, Europe, and the US? What is the incidence/prevalence? How is it treated today? Duration and cost of current therapy? What is in the pipeline? What are quasipublic hints of other development efforts? Who is in clinical trials? What about reimbursement and pricing? Are there regulatory guidelines and precedents? This requires far more than daily searching of the internet and publications or presentations. Sometimes a trip to a meeting is warranted just to confront a speaker in the hallway and ask the key confirming question: *Could I get a small sample of your new IL38 agonist?* Sharing information is key. Should this be on electronic bulletin boards accessible by others in the enterprise? Experts may assist in defining the targets and marketing support is essential.

Team concerns: Each project is unique but some questions are likely. Below is a brief outline of the first concerns by each member.

Chemistry: Is GMP material available? Cost? Purity? Is the manufacturer secure through launch? Will they improve the process—what is the slope of the learning curve? Will they keep the process secret? Are there generic suppliers? Is deuterated material available? Is an analytical method qualified for bulk drug, feed, and for human biological samples? Is the formulation adequate to begin trials? Is freeze-dried material stable? Should formulation enhancement be done intramurally or extramurally? What routes should be considered? Are toxicology formulations adequate for both acute and long-term dosing? Is clinical fill-finish a problem for initial and pivotal trials?

Biology: Are there animal models that are predictive of human biology? What is the minimal animal subset to satisfy the first human, his investigator, his IRB, FDA? Will a

Culex experiment with mice, rats, and guinea-pigs over one week provide sufficient preclinical pharmacology and kinetics? Will one-day toxicology suffice for the first human dose? How much genotoxicity and reproductive toxicology is needed for the first human? Are two-year studies going to be required? Can one be done after launch? Do you need 2-week, 1-month, 3-month, 6-month, 12-month toxicology in what species? What can you skip?

Kinetics: Human blood and *ex vivo* human microsomes can be used to characterize protein association and metabolism. Renal excretion often follows that in animals. Absorption from the mouth, stomach, duodenum, and jejunum may be tested in animals but also in the first human with a thin plastic nasogastric *feeding* tube. How much toxicology is needed for one human dose by nasal insufflation, inhalation, enema, suppository, or transdermal? When the first human is dosed, can a few samples be analyzed immediately to guide the dosing of the second human? One day's dosing in humans could be a small intravenous infusion of deuterated drug at the start with analysis of some samples and a larger oral unlabeled dose at 23 hours—absolute bioavailability in one day!

Clinical: Should the first human be dosed under an IND or in Canada or Europe? Just because you are in the U.S. doesn't mean that is the optimal site and FDA has approved drugs with no U.S. studies at all. Ensure ethics by using a super IRB to approve all protocols. Should the first humans be patients promised entrée to continuing studies or paid volunteers? After you have defined tolerance and kinetics in the first humans, can you proceed to pivots without Phase II? What biomarkers will help stratify?

Statistics–mathematics: Can you build **in silico** models of kinetics and dynamics and what do they predict about clinical dosage regimens? What are the minimal optimal observation points to confirm predictions of kinetics and dynamics? What are appropriate study designs that will fulfill regulatory expectations worldwide? Is there a crisp definition of the primary efficacy variable? Are the measurement scales appropriate? Is data collection crisp and self-validating? Can Bayesian approaches be used in sizing studies, interim analyses, adjusting size to achieve desired power, etc.? Would numerical permutation analysis be superior to traditional statistical parametric or nonparametric tests¹¹? Building a numerical model of key studies during the design phase, tweaking it with initial data, refining it as data flow in, and using it for final analysis is a powerful tool that can be shared with all team members.

Regulatory: Constant dialogue with regulators, in all the major regions, is essential. Consultants may help. There may be guidelines. Precedents may be uncovered. The FDA

and ICH web sites are useful. Plan for both regulatory success and blocks. A parallel IND strategy, for example, considers both Europe and the U.S. If the U.S. applies a clinical hold that is unreasonable, just do Phase I and II in Europe and then come to FDA with clinical data—the mice become irrelevant. Ask for every advantage—fast track, priority, rolling submission, putting some safety data in the safety update, etc. Is your product an orphan? That may help in Europe and the U.S.

Marketing: Marketing should be integrated with discovery strategies as well as development, but must be scientifically sophisticated¹². They should be challenged to produce real data on historical and current markets, reimbursements, pricing, etc. The uptake of sales, whether convex or concave, may make all the difference between profit and loss. The early addition of Europe and Japan may also make this difference. Marketing often wishes to try to *niche* market it in the U.S. with a small dedicated sales force, but should consider what the maximum sales will be, the time to reach peak, and the shape of the uptake curve. A larger more experienced sales partner may be able to double or triple sales. A European or Japanese partner, early on to help plan the study design, may be essential.

Legal: Intellectual property is essential and requires constant attention. Applications are expensive, require diligent pursuit, and must be global. Competition must be assessed constantly. A few patents are not sufficient—one should build an impenetrable estate. There may also be legal considerations early in planning when the label is written which will restrict marketing. Marketing may have a major input here as they must work with the label. Legal should also participate in informed consents and investigator contracts. Investigators should own their data and be able to publish it *ad lib* but with a copy of the submission to the sponsor. The sponsor also should register all clinical trials and make available summaries of their outcomes.

Family: A project team is a family. Each member has her own personality, but the family commitment to the project and its strategy must dominate¹³. Each member has good days and bad days and the family compensates and soothes. Each member must represent and garner resources from his *discipline*. The members are driven by tough timelines that may not be the highest priority for their colleagues in the *components* doing the actual work. That requires compromise and motivation and negotiation with the management of the *component*. A crisp definition of the strategy and the timelines helps such negotiations.

Definition of the goal: The next task is for the team to write the label proposed for their product at launch. What are the indications as each must be supported with adequate and well-controlled trials and limit what marketing can say. What are the dosage regimens?

What are the patient descriptors? What are the expected safety concerns? This will be the outline of all subsequent studies which should be directed at first proving that the product cannot achieve a required characteristics (**fail fast**)¹⁴ and then that it can be proved to have them. Thus if the dosing must be once daily, you perform kinetic analyses to find out if the duration of molecular presence or effect will achieve that goal. If not—stop. Careful *in silico* modeling is helpful at this stage. Note that *decision* research, before proof of principle and the decision to proceed to pivotal trials, is oriented to failing fast, but there is a transition at that point to proving safety and efficacy. This is a good point for transition of the team leadership to perhaps a marketing leader who will coordinate as well the global launches.

Preclinical efficiency: A crucial distinction is between academic publication and therapeutics development. We are all trained to replicate, replicate, control, control to get published and promoted. Development though requires that we restrain this impulse and do just enough to enhance the quality of the next decision. Two mice may be sufficient. Instead of hundreds of inbred identical mice with the same age, diets, genders, and genes, wouldn't human biology be predicted better by a few animals of different species or strains, different ages, different levels of obesity, etc.¹⁵? It may be promotable to publish the definitive study of the kinetics of a molecule in a thousand identical mice, but if it fails to be a drug who will read the paper? A new technology has revolutionized preclinical biology. The Culex¹⁶ encloses four rodents in isolated homes. Each has indwelling catheters in arteries, veins, the gut (e.g. gallbladder), etc. including microdialysis catheters, thermocouples, and ECG and EEG leads. Their activities are quantified but they are undisturbed. This is critical as entering the room occupied by a guinea pig can evoke a heart rate of 600 and alter drug kinetics and dynamics. What does gavage or blood sampling from the eyes do? The rodents have blood and excreta sampled periodically so that a single animal will define kinetics and often dynamics precisely. If four are studied each day, mice, rats, and guinea pigs can describe preclinical kinetics and dynamics in one week at four dose levels which is sufficient to define a smooth dose–response curve!

Work–out: Jack Welch championed work–out—simply not doing unnecessary stuff. At Eli Lilly and Company we had huge awards for Mickey Mouse achievement—employees identified unproductive procedures to be eliminated. I and other executives wore the hat and sang the song on stage while presenting the awards that were prominent in annual evaluations. As one can justify endless good scientific studies, that must be resisted. Yes it would be more precise to study hundreds of mice, but could two or ten suffice to provide enough insight for the pending decision? Yes the kinetics in ferrets might be interesting, but why not wait until you are sure the product will be approved. Any proposed study that is not linked to a statement in the label should be carefully evaluated. There is yet another very

useful work-out tool—failure aversion.

Failure aversion: To evoke critical discussions of studies list all studies required between team formation and the end result (which might be proof of principle or licensing rather than launch). For each of these assign the probability of failure that will be averted if the study is successful. For example, if the human kinetics are appropriate, that may be thought to reduce by 0.1 the overall probability of failure. At first you will find that these probabilities sum to a large number, as each proponent believes that her study is the most important. But they must sum to the difference between the probability of success and 1.0. The discussion in adjusting them to fit this sum is highly illuminating as each participant argues for his favorites and against the competing studies. When complete, you find that some studies have large scores and high priority. Others have very small scores or even zero. If a project when successful will have no impact on failure, why do it? It might be prerequisite to another important study, but often these are the historical feel-good studies every scientist will want to have but may not need. Ruthlessly challenge them¹⁷⁻¹⁸. Always ask if the project will be terminated if the study fails—if not, why do it? The ratio of this failure aversion score to the cost of the study (time, resources, risk, \$) is Ron Howard's¹⁹ index of priority—try to do the highest ratio studies first.

Project Map: Planning and modeling is a key team activity. One of the most important tools, and a graphic for the wall of the common room, is the map of the project²⁰. Map time on the abscissa (X axis) with days and real dates. As development is a series of critical decisions, identify them and mark their positions with vertical lines. These may include: first in vivo toxicology dose, first human dose, first dose to define efficacy, etc. Some studies must be completed so the results are at least partially available before a decision. Others may span a decision or must start afterwards. Map each proposed study as an oval with appropriate initial and final points. Within the oval express the failure aversion score and an estimate of cost (which may be crude). This is a critical exercise as each study will require that certain prerequisites be accomplished. Before the first mouse is dosed there must be GMP bulk drug, feed formulation, analytical methods, and a decision about the dosing regimen. Soon each of these milestones will be defined in days and the focus will be on whether the feed analytic method and proof of stability can be accomplished by day 127.

Timing: Project excellence demands quality-speed—no unnecessary detours, no delays, no waiting for a decision, careful planning, parallel not series, and a constant focus on minutes. As the Earl of Stanhope said: *Focus on the minutes and the years will take care of themselves*. On the critical path to a B\$1 product, each second costs \$35! Think that way. You can't schedule a meeting—decide now. You can't wait to get together—use

videoconferencing. Delegate. Perform. A realistic goal for the average total time from team to first approval is about 2000 days. In 1990 I evolved the **Win the Lilly 2500** program—2500 days from team to effective launch in 2/3 of available world markets. There were six unambiguous milestones (like first human dose) and each segment had a target number of days that had been beaten by at least one historical Lilly project. Overall this is 2000 days to the first approval and current goals are more aggressive. Guilford Pharmaceutical, for example, fashioned the **BtoB1000** goal²¹ of team to proof of principle in 1000 days, and they have beaten that. Inspire Pharmaceuticals is even more aggressive and successful. If everyone views quality–speed as essential, and works together to achieve it, project excellence can be achieved. But it takes a complete abhorrence of delay—no I won't call back, no we need the shipment tomorrow, no figure out a way to get my materials here by Friday!

Data: Paper is good for toilets—electrons are good for data and communication. Lilly now has paperless research, files, and regulatory submissions. FDA is promoting paperless research. Can you capture clinical data directly from patients? PDA diaries? Telephones, internet, videoconferencing? Why ask a professor of psychiatry to record the bowel habits of her patient? Video interrogation captures cognitive function, memory, history, adverse events, symptoms, etc. Paper is very expensive. It requires duplicates and corrections with an audit trail. It requires storage—warehouses of paper for decades. It delays analysis. It is expensive to validate. Why not capture the patient's data directly from the patient electronically²²⁻²³? Fill those filing cabinets with plants.

Reports: The team should define in advance all of the studies expected and all the reports that will be needed for global submissions. After it is clear that you will likely submit an NDA, it is appropriate to write in advance the key reports using blinded but real data. The SAS procedures can be finalized and the tables, graphics, and text can be prepared with what you expect the results to be. When the actual therapy assignments are known, the final tables and graphics can be prepared and the text edited. If you have predicted well, and are lucky, editing is straight–forward and this goes quickly. A two–week NDA (after the last patient visit) is possible²⁴. If you modularize the submissions, according to the ICH technical document formats, you can assemble modules for submissions in several regions with only a few local additions.

Stratification: *Pre hoc* stratification is key. Randomization is used to distribute evenly all variables except those controlled in the study, but it is not guaranteed. We hear of studies that failed because all the sickest patients were allocated to drug instead of placebo. We hear that the overall study failed but the drug was successful in women, by *post hoc* analysis. To ensure randomization, define all the variables you can measure that might influence the

primary efficacy variable (or safety) and stratify on them. If that creates some small strata there is no obligation to analyze every stratum, but it will prevent grossly uneven randomization that might sink your Bismark. At best, if the overall group fails, it still allows a legitimate predefined subgroup analysis.

Homogenous or realistic: If you could reliably identify a subgroup, such as young women, who would have the best responses to your new treatment, some argue that clinical trials should be restricted to that group to minimize the number of patients necessary to achieve the regulatory hurdle of $P < .05$. But if you aren't confident of defining such a subgroup and if you are going to market to a more diverse population, isn't it wiser to include in trials the more diverse group, simulating the market. Stratification allows you to have it both ways—you get a small subgroup you can analyze to prove efficacy and safety at least in that group and you predict what the market will look like²⁵. As safety concerns in off-label use can damage your product, this is much safer. If a problem will arise in elderly women also taking lipid lowering drugs, best to discover that in a trial where they are protected and when the signal can be discerned from a few patients than to have newspaper headlines of the slaughter of innocents.

Interim analyses: Scientists prefer double-blinding to avoid patient or observer bias. Regulators, concerned that sponsors may cheat on patient enrollment or data interpretation, often insist on triple blinding. This does not preclude interim analyses. A study may be designed such that at critical points a predefined blinded analysis will be performed by the computer program that contains all the data and randomization codes. If a dose is performing badly, it can be eliminated or fewer patients assigned. If a serious problem is detected, a flag may be raised signalling the need to unblind the data like an onion, preserving blinding of individual patient treatment allocation if possible²⁶. Such careful planning, though seemingly complex, can be highly efficient. In 1984 FDA and European regulators agreed to a single clinical protocol for quinelorane after first human dosing. This D2 agonist began trials in 1024 patients with four doses, two genders, and three illnesses (sexual dysfunctions) for 24 cells but the protocol allowed for adding three doses and two more illnesses (or eliminating those that didn't work). This single protocol was going to be divided statistically into *two* and serve for all pivotal studies. Because of nausea that couldn't be eliminated, marketing/sales killed the project but it worked ($P < .05$) safely in two of the predefined illnesses at low doses—the first agent proven effective for sexual dysfunction in both men and women.

Work-out Phase II: Phase II is where sponsors who lack profound knowledge or effective teams flounder around trying to figure out what their product might do. Instead,

initial human dosing should define the maximum safe dosing regimen for pivotal studies. Begin a dose–response pivotal study with that dose and one or two smaller ones, as regulators will require you to define a no–effect or at least less–effective dose. One dose rarely is optimal for all patients. Design appropriate interim analyses. Alternatively, use adaptive allocation to continuously alter the probability of assignment to doses according to preliminary estimates of their effects. If you find you have chosen wrong and the study doesn't work, you have then completed a Phase II study. If some of your doses are safe and effective, you have skipped Phase II and probably one or two years of floundering.

Project evaluation: You need a way of expressing succinctly the strategic value of your project as well as tracking its progress and comparing it with other projects. There are four critical variables: cost, value, time, and probability. Defining them leads to excellent discussions among experts and can draw out key concerns from those reticent to speak up. This is how profound knowledge is acquired. Each variable may be discussed by the team plus others with expertise to contribute to these discussions.

Cost: Cost is the least important variable. For most projects it is not that critical (time being more valuable), although a lack of unlimited resources certainly drives biotechnology strategies more than those of big pharma. Costs can be estimated sufficiently from historical models stage by stage and adjusted for special concerns. There are proprietary databases, such as CROCUS, that are good indicators of clinical trial costs and most vendors will estimate preclinical costs.

Value: This is the most important but also the most difficult of the variables to quantify. Competition is a key determinant but good intelligence on competitive projects is illusory. Actual historical market data are available and should be scrutinized to examine each country's uptake curves and the market shares of current therapies. Pricing and reimbursement are becoming more problematic. Intellectual property protection may become dominant in estimating value. When the team is first formed, it may be sufficient to use a point estimate drawn from historical data on similar products, but as the project advances more profound knowledge should refine such estimates.

Time: This should be considered in days and real dates—not months and half–years. Start with each critical decision and each unambiguous milestone. What are historical records? What are the quality–speed goals? Initially it may be sufficient to focus on the timing of U.S. launch, but Europe and Japan must be considered as today the investment in development must be repaid by returns from all three continents.

Probability: This is the best way of following the progress of the project. When the team is formed the probability of successful launch is probably 0.05 to 0.1. As studies are completed and more profound knowledge garnered, the probability should rise. It should be 0.8 or more at proof of principle before major pivotal studies. The *expected value* is the product of the probability and value if launched and so you are generating expected value. The difference between the cost and the expected value generated is a measure of the *bang for the buck* of this project. If the probability isn't ascending rapidly, the team has got a loser and it is may be time to fail fast.

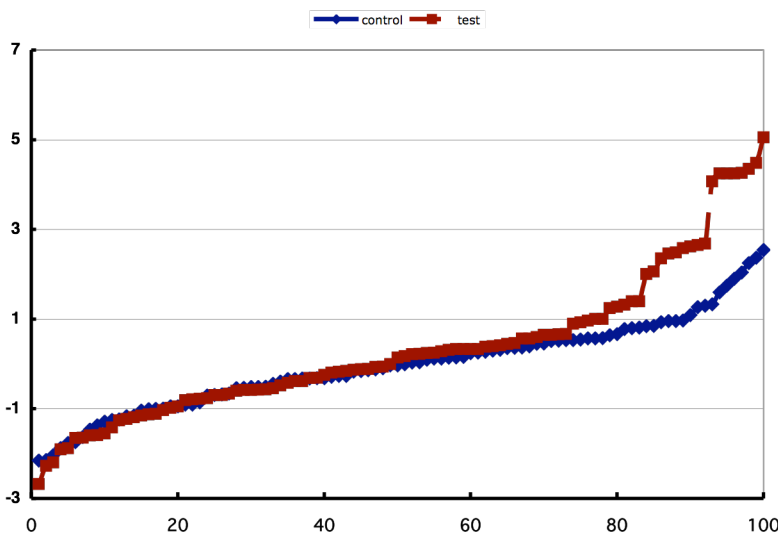
Portfolio: If each project is defined in terms of cost, value, time, and probability, then the portfolio is easily displayed as these four variables²⁷⁻²⁸. Often a VvsP and a CvsT pair of plots is easy to work from, although there are ways of displaying four dimensions (a 3-D display with cost being different sizes of the icon)²⁹. Colors can be used to display other strategic values such as therapeutic or pharmacologic platform. Usually when each project is well-defined and everyone agrees on the validity of those characterizations, managing the portfolio is both transparent and relatively simple.

Strategic coherence: Of critical importance in a biotechnology enterprise is coherence of strategies from discovery through marketing/sales worldwide³⁰. It is a disaster to have discovery resources devoted to something that will never fit with the development, regulatory, or marketing/sales strategies and cannot be quickly licensed out at a profit. Development resources must be carefully managed as although much can be rented, it still requires intramural expertise to plan, monitor, and coordinate the process. Big pharma has muscle that may be useful in global pivotal studies, global regulatory adjustments, and global marketing/sales. Often biotechnology firms should invest in generating value through proof of principle with an expectation that they can then profitably license their discovery.

Quality-speed: Operational excellence is achieved through quality-speed³¹. It is a form of Total Quality Management. One prospectively defines all those elements that may cause problems, and builds protections from failures. One eschews any delays and rigorously avoids and corrects errors. A six-sigma attitude, striving for less than one error per million, is a good goal and the current rate of correcting clinical data on paper clinical report forms greatly exceeds that. Don't rewrite—write once, use often. Define formats and labels in advance. Plan for studies, data, reports, tables, graphics, etc. and outline for them and then drop them in place when complete. Remember the Olympic athlete—one hundredth of a second faster without fouling and you win immortality instead of obscurity. That hundredth of a second takes thousands of hours of training, planning, and preparation!

Study and project goals: In addition to the obvious primary goal of both, each should have for the enterprise a defined second goal—to improve the process. Each study should set new quality–speed records and define how to better them on the next study. Each project should be meticulously autopsied with records of each decision, how it could have been done earlier, and how a similar project started today would be better performed. Only through such continued improvement will the enterprise leverage its existence to achieve more than a newcomer. Big pharma has failed to accomplish these secondary goals, but biotechnology enterprises are often neglecting them. You can learn from others in your industry and other industries, but constant improvement, a constant quest for knowledge, and a refusal to accept second best or slow progress is the ticket to success.

Data inspection versus analysis: R. A. Fisher, the father of much of modern scientific statistical analysis, emphasized the need to inspect the data before applying any tests. A good beginning is to array all the data sorted from smallest to largest in a probability



distribution as shown below.

These are data from 100 members in each of two groups treated differently. The smaller values are normally distributed and overlap. At the higher values there is obvious divergence. R. A. Fisher also emphasized permutation tests. The vaunted but flawed P value is a measure of how unusual is the distribution of data you observed. The best measure is to use all the data you observed, but create new samples by randomly allocating each

datum to one of your treatment groups. If you are measuring the difference as some metric of central tendency (e.g. mean or median), then determine that for each of your artificial samples. Array all your samples, perhaps 10,000 of them, from smallest to largest metric and just count how many had values of the metric as great or greater than the original sample of real data. The proportion of your randomized samples more extreme than you observed is the real absolute P value that is not dependent on data distribution. In doing this analysis you can use all your strata and make the samples as complex as you like. In the example shown, a two–sample t test not assuming equal variances, that assumes the data are normally distribute throughout, has a $P > .05$ but permutation testing yields $P < .05$ —the difference between bankruptcy and approval to market. More important, consider the

implications of a set of patients with clearly better response.

People not herds: To protect the health of the general population, regulators require proofs of safety and efficacy in the study population or a predefined subgroup. This assumes that use of the product would be random throughout the patient population, but that is unrealistic. In practice, one tries a new therapy and if it works you continue and if it doesn't you adjust the dosage regimen or quit. If there is a subgroup that has good responses, you attempt to identify it through measured characteristics. But, if you can't, you can define the timing of tests to distinguish such a good response so that the user can then decide whether to continue. To establish such a selective subgroup advantage may require a large or long study that the sponsor can't afford. To address this issue the FDA has used Subpart H approvals for drugs that have sufficient proof of general safety and good evidence that efficacy is likely, allowing them to be sold but with a defined requirement of completing an agreed additional protocol to establish efficacy more definitively within a time certain. For serious illness without alternative good treatments, such conditional approval is rational and humane. As long as safety concerns are collected rigorously and reported promptly, this would do little harm and may afford access to valuable treatments to those who might otherwise die before traditional approvals.

Differential licensing: At present all physicians are licensed for all medicine. Should new unestablished therapies be restricted to physicians who will rigorously observe safety signals, follow protocols, collect valid data, and communicate well with sponsors? Lou Lasagna suggested three regulatory actions—a disapproval red light, a general approval green light, and a selective sale but reporting of use as an orange caution light. I believe that HMOs have an opportunity for early introduction of new treatments in return for rigorous adherence to a protocol with good reporting of results. That would allow naturalistic large trials to be done economically and quickly.

Biomarkers: There is fond hope that biological tests will discriminate patients who will respond best to a treatment, identify those with special safety concerns, and facilitate dosage optimization. Although the genomic tests are new, the concept is old and was the rationale for the emergence of clinical pharmacology five decades ago. Early biomarkers include measurement of acetylator phenotype (to adjust doses of drugs such as INH), antimicrobial sensitivity testing (to choose drugs), and recognition of the influences of gender, age, smoking, charcoal grilling, etc. Some important conditions remain refractory—Prozac 20 mg/day works in 2/3 of depressed patients, and no other dose is better and no test identifies responsive from unresponsive patients. The problem with pharmacogenomics³² is that it requires large clinical trials to identify the marker, the tests are not readily available

and are expensive, and general physicians do not today use such information—they don't even use obvious biomarkers such as gender and age when the implications are well known. The vision of a drop of your blood replacing physician knowledge and judgment is unlikely.

Centers of Excellence: NIH could now fund centers to study, teach, and implement new methods of critical path research and FDA could aid such efforts. Now is the time.

References

- ¹ Stuart H. Yuspa MD; yupas@de37a.nci.nih.gov; nihroadmap.nih.gov/overview/asp
- ² FDA. Innovation Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products. 2004 March
- ³ Thompson WL. Therapeutics in the third millennium: Strategy and tactics to achieve the apogee of value creation. *Profound Quality Resources Monograph* 1998Oct; 34:1-40.
- ⁴ Profile Pharmaceutical Industry 2004; www.phrma.com;
- ⁵ Windhover's In Vivo: The Business & Medicine Report: Bain drug economics model, 2003.
- ⁶ Thompson WL. Why do drugs cost so much in the U.S.? *Lilly* 1994Feb;1-10.
- ⁷ Thompson WL. Fractal systems re-engineering lubricates innovation to enhance rusty rigid feudal structures that inhibit function. *PQR* 2004Jun;90:1-12.
- ⁸ Thompson WL. Thriving in the Third Millennium: Quality-Speed strategies for therapeutics research and biotechnology liaisons. *Profound Quality Resources Monograph* 2000Jan;37:1-68.
- ⁹ The Management Premium—The Link between Operational Excellence and Shareholder Value in Biotech Firms. *Biotechnology Management Practices* 2003. Mardis Aibel & Associates, LLC. The Business and Medicine Report, www.windhover.com.
- ¹⁰ Thompson WL. Exhilarating Project Management for that thrilling roller-coaster ride to enhanced Quality-Speed-Cost. *Profound Quality Resources Monograph* 1997Mar; 29:1-16.
- ¹¹ Thompson WL. Contemporary design and analysis of clinical trials. *PQR Monograph* 2002Apr;44:1-15.
- ¹² Thompson WL. Mastery of quality-speed-cost of drug development through excellence of decision research. *Profound Quality Resources Monograph* 1998Jan;32: 1-8.
- ¹³ Thompson WL. Quality—speed meetings. *Profound Quality Resources Monograph* 2004Jan; 82:1-4.
- ¹⁴ Thompson WL. Dix, cent, mille: Establish efficacy in 100 patients in 1000 days for \$10 million (Chapter 6, pp 93-111). In: Cato A, Sutton L, Cato A (eds). *Drug Trials and Tribulations*, 2ed. New York:Marcel Dekker;2002(ISBN 0-8247-0314-6).
- ¹⁵ Thompson WL. Therapeutics Renaissance trials 2004. *Profound Quality Resources Monograph* 2004Apr;86:1-8.
- ¹⁶ Kissinger CB, Kissinger PT. Can preclinical ADMET-PK now be done more efficiently and effectively? *Preclinica* 2004(Sep-Oct);2(5):319-323.
- ¹⁷ Thompson WL. Clinical trial tactics for the third millennium: Fast, Focused, Frugal & deFinitive. *Profound Quality Resources Monograph* 2000Dec; 41:1-56.
- ¹⁸ Thompson WL. E4 clinical trials for the third millennium: Ethical, eclectic, elegant, elucidating. *PQR Monograph* 2000Jul; 38:1-48.
- ¹⁹ Ron Howard, ScD, Professor & Chairman of Engineering Economic Systems, Stanford University; father of Decision Analysis.
- ²⁰ Thompson WL. Quality-Speed from Bench to Bedside II. *PQR Monograph* 2002 Sep; 57:1-12.
- ²¹ Thompson WL. Bench to bedside with quality-speed. *PQR Monograph* 2002 Jul;47:1-8.
- ²² Thompson WL. Health informatics. *Profound Quality Resources Monograph* 1995Apr; 16:1-9.
- ²³ Thompson WL. Electronic standards for transmission of regulatory information (business case; ICH). *PQR Monograph* 1995 Apr;17:1-7
- ²⁴ Thompson WL. Zen and the art of the two-week NDA. *PQR Monograph* 2003Sept; 74:1-6.
- ²⁵ Thompson WL. Patient-centric trials (PCT). *PQR Monograph* 2002Jul;48:1-9.
- ²⁶ Thompson WL. Interim Analyses: for ethics, frugality & speed. *Profound Quality Resources Mono.* 1999Jan; 35: 1-8.
- ²⁷ Thompson WL. Fractal portfolio resourcing. *Profound Quality Resources Monograph* 1995Jun; 21:1-44.
- ²⁸ Thompson WL. Seamless process re-engineering. *Profound Quality Resources Monograph* 1995Jul; 22:1-56.
- ²⁹ Thompson WL. Cost, time, value & probability portfolio analysis. *PQR* 1998Jun; 53:1-4.
- ³⁰ Thompson WL. Strategy. *Profound Quality Resources Monograph* 2004July;91:1-8.
- ³¹ Thompson WL. New Pharmaceutical Obstipation—Four Decisions as Effective as Exlax. *PQR* 2003Mar; 63:1-6.
- ³² Thompson WL. Informatics, genomics, and medical reformation—thriving in the third millennium. *Profound Quality Resources Monograph* 2000Nov;40:1-12.