CLINICAL INVESTIGATION

An evidence-based approach to conducting clinical trial feasibility assessments

Otis Johnson
inVentiv Health Clinical
otis.johnson@inventivhealth.com

Failure to enroll patients is now one of the leading causes of clinical trial delays – and even of trial failures. By using a data-driven approach to determining trial feasibility, drug sponsors can specify with a high degree of confidence exactly how long it will take to fulfill a study's patient quota. Today, there are rich sources of data that can inform the trial forecast, as well as sophisticated modeling tools to test 'what if' scenarios and establish confidence levels for the results. Using these resources, sponsors **can** accurately forecast enrollment timelines, flag potential risks to the plan and identify variables that can be adjusted to influence the timeline.

Keywords: clinical trial • enrollment forecasting • feasibility • patient enrollment • simulating enrollment

Sponsors of drugs in development and the Contract Research Organizations (CROs) working on their behalf are often caught off guard by the difficulties they experience in enrolling subjects into trials. Indeed, timelines for the majority of trials must be extended in order to fulfill their patient quotas. As the traditional approach to forecasting clinical trial timelines – relying on investigators' estimates – is failing sponsors, it is time to take a new approach. We propose a comprehensive methodology for researching recruitment feasibility that incorporates patient insights, draws on commercially available data and uses statistical modeling to improve the predictability of patient enrollment.

Patient enrollment statistics

The biopharmaceutical industry has long operated on the belief that enrolling patients in clinical trials is easy enough, provided that the protocol is right and that the best investigator sites have been enlisted. The facts, however, suggest that although this might have been the case in simpler times, it is no longer so. Consider that:

- 11% of selected sites *never* enroll a single patient [1];
- 48% of all sites underperform, meaning that they do not deliver the number of patients they expect [2];
- Sponsors' original timelines for Phase II–IV studies usually end up doubling in order to meet the desired enrollment levels [2];
- Nearly 80% of clinical trials fail to meet their enrollment timelines [3];
- Although results vary widely by therapeutic area, on average, only half of all patients screened complete clinical trials [2].

Many enrollment forecasts are faulty, causing sponsors to embark on trials with a false sense of what they can accomplish in a given time frame. The resulting deviations from enrollment forecasts have a significant negative effect on research costs and corporate revenues, and in many cases, the trial budget gets larger. The daily operational costs of running a trial are an estimated US$37K, and any timeline extension results in additional research costs that were not accounted for in the trial budget [4]. What

FUTURE SCIENCE part of fsg

is more, the opportunity costs of delayed market entry are an estimated $600K to $8M a day [5].

Traditional feasibility practices

The difficulty manufacturers have in bringing trials in on time and on budget is not for lack of planning. The fault lies in the fact that the plans are generally based on unreliable or partial information. Once they have developed a recruitment plan, sponsors attempt to confirm that their timelines are realistic. They routinely conduct 'feasibility exercises' that involve working with their CRO and sites (traditional research sites, site networks and integrated health systems) to answer the following types of questions:

- What regulatory hurdles can we expect to encounter in the countries being considered for the study?
- Is the study design consistent with the standard of care in the various regions?
- Is there a large enough patient population to justify inclusion of particular countries and sites?
- Are the sites able to get the right infrastructure and staff in place to support the trial?
- How many patients can a given site enroll in a specified time frame?

Sponsors, or their designated CROs, then process the enrollment information from the surveys to establish enrollment plans. Often, these plans are simply based on extrapolation of a mean enrollment rate to determine the number of sites needed and the expected enrollment duration. This approach does not consider various factors that can affect enrollment timelines. Chief among these often-overlooked factors are that 11% of sites that do not enroll any patients [1] and reliance on just the midpoint to make projections will skew the results.

In the end, the sponsor wants confirmation that its plan is realistic and will result in a successful trial. The CRO wants to be sure that it can comply with the sponsor's expectations, enrolling the trial within the deadline. And sites want to ensure that they can deliver quality results, and, through their contact with sponsors and CROs, want to learn of upcoming trials of potential interest.

Because most sponsors primarily rely on investigator sites to carry out trial recruitment, they, quite logically, assess recruitment feasibility by surveying sites on their projected enrollment capability. They ask for input via questionnaires on how many patients who fit the study criteria an investigator could expect to recruit for an upcoming trial. Companies then, typically, temper the results with their own intuition and judgment to estimate how long it will take to recruit the necessary number of target patients. This tempering is necessary because investigators tend to overestimate their enrollment capability, hoping they will appear more attractive to sponsors and CROs and get selected for the trial. How well this approach works depends on the experience of those involved. At times, it may produce a valid estimate, but the above statistics on enrollment performance suggest that it usually does not.

Limitations of relying on feedback from investigators

There is nothing wrong with surveying investigators to gather input into study feasibility; it is simply that this step alone is inadequate to produce accurate enrollment forecasts consistently. The reasons are many, some of which have to do with the validity of sites' responses, since:

- Busy physicians do not tend to run queries against a patient database to answer sponsors' questions about available patients with any precision; they simply provide a rough estimate;
- Physicians cannot divine the future, but at best can only make an educated guess as to how many patients they would be able to furnish;
- Sites tend to be overly optimistic and to overcommit when it comes to recruiting patients – as observed by Dr. Louis Lasagna many years ago and now dubbed 'Lasagna's Law.' This is not ascribed to any intent to deceive, but rather to do what is best for their patient. If the proposed study could be advantageous to their patients, investigators have a vested interest in estimating their enrollment potential on the high side so that they will be selected for participation (see Figure 1).

Other limitations stem from the way sponsors handle the information they get from sites. For instance:

- Sponsors routinely put too much credence in what investigators tell them. Experience suggests that investigator estimates are often inflated by as much as 75%;
- Often, sponsors do not dig deep enough in asking sites about their enrollment potential, or do not necessarily understand what sites take into account when providing their answers. One site may estimate its enrollment potential based on the total number of patients it sees in the therapeutic area while another

may consider how many patients are likely to meet the trial's specific inclusion/exclusion criteria;

- There is a tendency to assume (incorrectly) that all sites will recruit patients, beginning at the same time, and at the same rate. In actuality, this is quite variable, as there are differences in ethics committee review and approval timelines, differences in site initiation timelines by country, differences in access to patients and differences in allowable patient recruitment methods by region and country.

Conducting due diligence

A methodical, multistep process that uses all available information (from internal and external sources) can be used to predict the probability of enrollment success in a specific time frame, given certain variables, with a high degree of accuracy. The recommended approach incorporates information gleaned from physician surveys as only one piece of a much broader, more intense due diligence process that considers the patient perspective, taps commercially available databases and uses statistical modeling.

The ideal time to begin this work is during the preparation of the Clinical Development Plan (CDP). Thus, findings can be incorporated into the plan while there is still time to influence the protocol.

A comprehensive feasibility assessment involves gathering information as detailed below and then using statistical modeling to incorporate and manage it all. In reality, the following steps are performed in this general order, although not in quite as linear a fashion. The due-diligence stage of trial planning involves:

- Gauging patient availability;

- The first step in the process is to estimate the number of patients who will be eligible to participate in the trial. This calculation will be based on a breakdown, by country, of the incidence/prevalence of the disease and competing trials in the same therapeutic area; the study inclusion/exclusion criteria; the treatment guidelines and procedures;

- Electronic Medical Records (EMRs), prescription, and/or integrated medical claims databases should be mined to understand how many patients are being treated for the particular disease and where they are located (see Figure 2). These databases can be searched by diagnostic and procedure codes, as well as by various other inclusion/exclusion criteria. One caveat is that the information contained in these databases is not always sufficient to evaluate patient counts based on a study's particular inclusion/exclusion criteria. For instance, it may not be possible to tell from EMRs a patient's predicted FEV_1 (forced expiratory volume in 1 second – a lung function measurement), as this information is often captured on a separate device and may not get uploaded to the central EMR system;

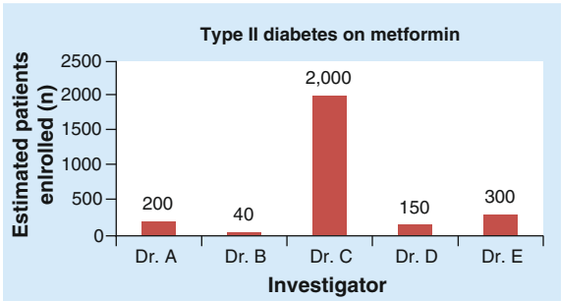

Figure 1. Sample investigator estimates of patients with Type II diabetes, taking metformin. Five investigator sites reported the number of patients that they believed fit the eligibility criteria for a trial for an antidiabetes treatment requiring that patients be stabilized on metformin. Their estimates vary widely, and taking an average of these estimates (538 patients per site) will be very misleading to study planners. A more realistic approach is to use the median of 200, or to disregard the 2000 as a potential outlier if additional follow-up information cannot be obtained from the investigator reporting this information. The enrollment data set should then be processed to examine the median (175) and 25th to 75th percentile range (68–275) to develop a good understanding of patient access across the investigators.

- Gathering the patient perspective;

- At an early stage in the trial planning process, it is important to understand how the target patient population would respond to the protocol requirements. What would patients find appealing or objectionable about the approach? What motivational drivers would influence their participation decision? It is possible, for instance, that some aspect of the protocol could pose an unforeseen emotional or logistical hurdle for patients and be a risk to enrollment. Taking the time to assess this is in complete harmony with the US FDA's 'Patient-Focused Drug Development' initiative that 'aims to more systematically gather patients' perspectives on their condition' [6];

- A combination of primary research and social media monitoring should be used to validate assumptions made about patient attitudes, beliefs and behaviors so that the protocol can be evaluated through a patient's lens. When discovered early, any possible

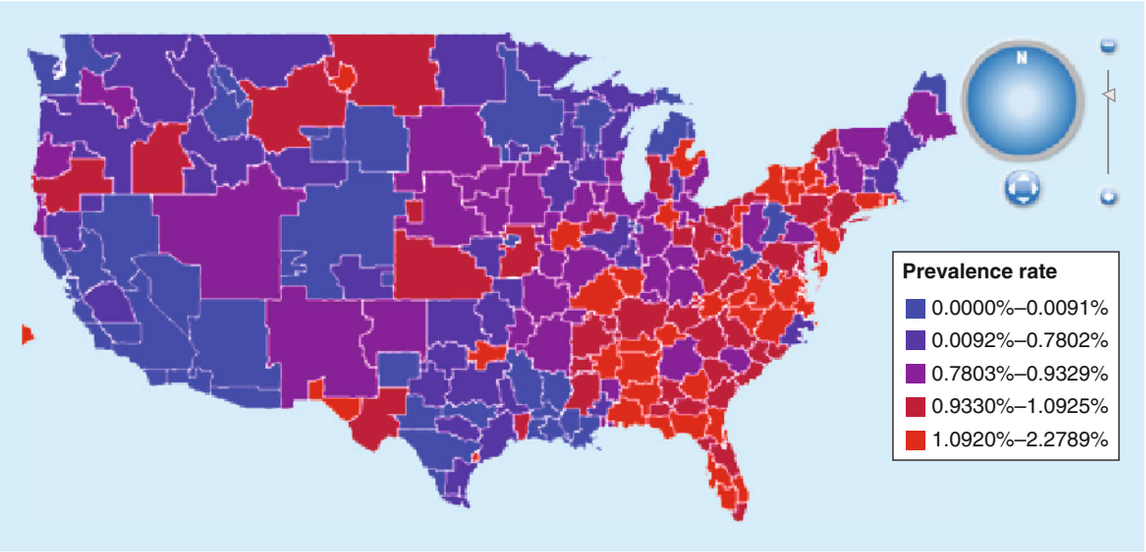

Figure 2. Incidence of Type I diabetes excluding patients with chronic obstructive pulmonary disease. Illustrates the incidence of patients in the target population (Type I diabetes, excluding patients with chronic obstructive pulmonary disease), by location.
Reproduced with permission from [8].
For color images please see online: http://www.future-science.com/doi/full/10.4155/CLI.14.139

risks can either be mitigated or factored into enrollment estimates. For example, in the early stages of planning one study in overactive bladder, patient feedback had a significant influence on the way the sponsor proceeded. In a focus group, patients stated very strongly that they would not subject themselves to an ophthalmology procedure that required them to sit in a dark room for 30 min. The protocol was actually canceled based on this feedback;

- Selecting countries with the most potential. The next logical step is to determine the optimal mix of countries based on their potential for enrolling patients. This is a matter of blending the information gathered above on patient counts, analogous historical trials and competing trials with information on where sites operate by therapeutic area, indication and specialty. A number of commercially available databases can reveal how many sites have the relevant experience, by country;

- The number of available patients can be affected by the number of competing trials in the same indication in a given country. The global, competitive landscape can be surveyed using a combination of clinical trial registries, publications and subscription databases to identify areas where competing trials are/are not likely to make recruiting difficult, while targeting areas with an adequate number of experienced investigators. Another valuable piece of intelligence can be gained by gathering intelligence on comparable historical trials. Visualizations, such as the ones below in Figures 3 & 4 can illustrate for trial planners where sites meeting basic criteria are most concentrated [7]. These visualizations and other supporting analyses are then documented in a country-level decision matrix similar to Table 1 to facilitate selection of the most appropriate countries for a particular trial;

- Identifying the best investigators/sites. A major factor in enrollment success is site selection. Companies tend to turn repeatedly to a limited number of sites with which they have prior experience – an approach that often leads to disappointing results. A better approach is to use research and analytics to target the right sites that have:

  - The ability to enroll the most patients, given their proximity to patients;

  - Determining sites' proximity to patients is a matter of overlaying a database of clinical investigator listings (such as the FDA's 1572 database in the USA) onto a patient heat map, similar to that depicted in Figure 2. The result identifies investigators who are colocated with concentrations of patients who meet the study criteria, making them prime targets, at least in terms of their access to patients;

  - A history of strong performance, both in terms of recruitment and data quality;

- Evaluating sites based on their prior performance can involve, at the most basic level, reviewing historical data on analogous products from within a company's clinical trial management system. It should also include reviewing public sources such as www.clinicaltrials.gov. As internal sources are usually limited in scope and public sources are limited in detail, the ideal solution is to tap commercial databases maintained expressly for this purpose as well. Through such a service, investigators are assigned an objective, composite performance score based on the number of trials they have participated in and their past enrollment performance, to include initiation periods, screening rates, and failure rates;

- The required capabilities in terms of staff, experience, infrastructure and equipment; Most protocols will require that sites have specific capabilities, whether it be expertise in the therapeutic area, familiarity with a particular technology or access to specialized equipment. Analytical tools are available to profile sites on a range of dimensions, including their research activity, infrastructure, personnel, initiation timelines and access to the targeted patient population;

- Sufficient capacity without competing trials;

- This can best be established by using the technique that sponsors have traditionally relied on to the exclusion of most others: sending questionnaires to sites. Once sites have been identified that appear to meet the above criteria, they can be surveyed as to their availability in the proposed time period. Whether they are participating in competing trials can also be verified though publicly available and subscription trial intelligence databases.

Note that conducting this research by mining available databases does not preclude the need to survey sites. However, it does mean that the questionnaires can be much shorter, reducing the burden on site personnel. It also is a much more evidence-based approach that overcomes the inherent bias in site-reported information and considers sites' applicability across multiple factors that affect performance.

Simulating enrollment rates

The findings from the above research should then be converted to quantitative inputs that are run through an assumption-based enrollment model as a way of determining if the right planning parameters are in place for the trial. The specific statistical technique, Monte Carlo Simulation, calculates the probability that a particular outcome will occur based on a given action or set of assumptions. It considers the selected country and site distribution and works by assigning a range of values to each input variable – such as screen failure and site failure expectations, patient availability, enrollment performance and the effect of patient outreach campaigns – and then running a series of virtual trial simulations. Typically, the software will run 1000

Figure 3. Investigator availability at a glance.
Investigator availability for Type II diabetes mellitus from a site intelligence database. This suggests that ideal locations for the study based on investigator access (one of several country selection parameters) include parts of Asia, South America, Europe and the USA, given the clustering of sites experienced in this specific indication.
Reproduced with permission from [7].

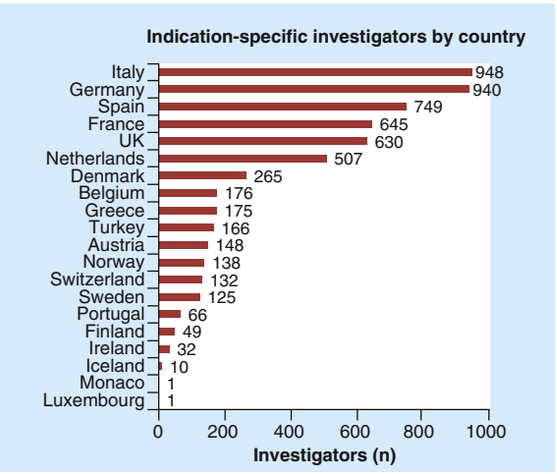

Figure 4. Number of indication-specific investigators by country. This figure shows availability of investigators experienced in the indication by country, a key data point in country selection decisions. It can be used to select the European countries best suited for the trial being planned.

Table 1. Country decision matrix[†].

Country	Relevant trial experience	Competing trials	Company experience	Site access	Patient access	Initiation time line	Enrollment potential	Assessed score
USA								
Russia								
UK								
Canada								
Germany								
Mexico								

[†]Populate with data from due diligence reaseach of internal and external data sources.

to 5000 simulations, with each using a different set of random values from within the range.

This type of assumption-based modeling is especially useful for program planning, as multiple scenarios can be modeled. The results are then displayed as a distribution chart showing the probability of meeting enrollment targets for each scenario (see Figure 5).

This level of precision in enrollment forecasting has never been possible before and stands to improve the trial enrollment process dramatically by setting realistic expectations from the outset. This model is now in routine use, and we are monitoring studies going through the process and will have results to report in a year.

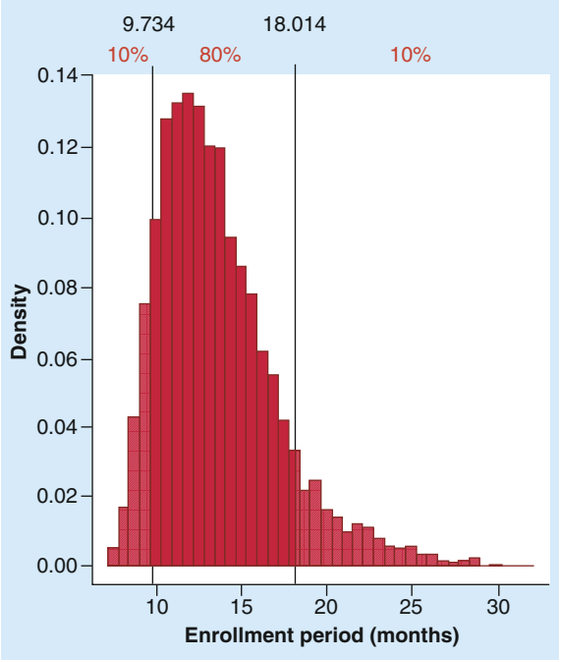

Figure 5. Enrollment probability by month. A Monte Carlo Simulation can be used to model the probability of enrolling a trial within a given time period. The output of this particular simulation reveals that with a given set of variables (specific countries and sites), there is only a 10% probability of enrolling the trial in 9.7 months and a 90% probability of doing so in 18 months. If the goal is to enroll the trial in 10 months, certain variables will have to change for there to be a strong probability of success.

Case study: retrospective analysis

A leading global pharmaceutical company undertaking a Phase II clinical trial of an osteoarthritis therapy engaged its CRO to randomize 200 patients into the trial. The pre-determined plan called for completing patient enrollment in 8 months, using 25 sites that the sponsor had selected based on its past experience with them.

After several months of sluggish enrollment, the sponsor decided to implement its contingency plan and added eight rescue sites. The maneuver was both expensive and time consuming. Activating each additional site cost $160K, plus monthly maintenance costs of $1.5K per site. Despite the money spent, recruitment remained suboptimal. It ultimately took 11.5 months to enroll all 200 patients, 3.5 months longer than planned. Of the 33 total sites, four had recruited no patients.

Although the sponsor did not have the benefit of planning its trial with the information and analytics proposed here, it is possible to demonstrate retrospectively how doing so might have changed the outcome. When the original parameters for the trial were put through the simulation model described above, the results explained why the trial had not gone as planned: there had been only a 20% probability of success in using the selected 25 sites to recruit 200 patients in 8 months.

To test the power of the evidence-based approach to site selection, analysts, retrospectively – and without knowledge of the sponsor's original list of sites – produced a fresh list of recommended sites. They mined the databases described above and ranked sites based on their proximity to target patients and objective ratings of their past performance. Only 15 of the 33 sites that the sponsor had involved in the study were among the top sites that surfaces as a result of this research.

These 15 sites had a proven history of randomizing three times as many patients with significantly fewer screen failures than those not selected.

The simulation model was run again, this time with better planning parameters and a list of sites drawn from the database research. The output from the simulation exercise indicated that the sponsor would have had an 85% probability of successfully enrolling the trial in 11.5 months using 29 sites selected through the process. This is how long it actually took the sponsor. Unfortunately, because the sponsor did not turn to the best sites, it had to involve 33 (rather than 29), and it missed its enrollment deadline by 3.5 months (Table 2).

Conservatively, this company spent close to one million dollars in unbudgeted funds to rescue this trial and, perhaps even more important, had to reforecast its revenue projections for the product, given its delay toward approval. Had the sponsor gone about the feasibility exercise in the way recommended here, the company would likely have been spared the expense, disappointment, and disruption of not meeting its goals.

Tips for success

Study planners wishing to apply an evidence-based approach to evaluating a study's feasibility should:

- Begin the feasibility assessment before the protocol is finalized. Ideally, it should begin during the creation of the CDP;
- Take a multidisciplinary approach to performing the feasibility assessment. The best teams include members with experience in Clinical Informatics and Analytics, Medical Affairs, the given therapeutic area and Clinical Operations;
- Work closely with the medical director aligned with the therapeutic area to ensure that they have a comprehensive understanding of the target patient population;
- Include patient insights early in the process to rule out any risk to enrollment stemming from patient attitudes, beliefs and behaviors;
- Rely on objective data drawn from sources other than sites themselves when assessing sites' enrollment potential;
- Gather information from sites as part of the process, but do not automatically take their enrollment estimates at face value. Delve deeper when asking about their patient population and factor their estimates into a broader analysis that incorporates data from several sources;
- Take advantage of statistical modeling to assess operational parameters and gain confidence in the chosen options;
- Avoid 'planning at the midpoint.' When sites respond to feasibility questionnaires with wildly different estimates, working with an average or mean is going to be misleading. Rather, it is more accurate to look at the distribution curve of responses, decide how to handle outliers and consider the full performance distribution in your planning process. This will produce a much more realistic idea of sites' enrollment potential;
- Flag potential risks during the feasibility study, and brainstorm ways to mitigate them;
- Use technology to the fullest and continue to evaluate emerging technology as new innovations and methodologies are always becoming available.

Table 2. Comparison of recruitment projections.

Scenario	Enrollment time line (months)	Number of sites	Probability of success (%)
Sponsor's original trial enrollment plan	8	25	20
Sponsor's actual results	11.5	33 (four of which recruited no patients)	
Sponsor's projected performance (calculated retrospectively) with evidence-based site selection	11.5	29	85

In a real-life exercise, a company's projected patient enrollment was modeled using Monte Carlo Simulation. The results indicated that the company had only a 20% chance of succeeding within its established time line, given the sites it had selected. In fact, the company did not meet its trial goals. It needed an extra 4 months and four rescue sites to complete enrollment. The same model indicated that the company would have had an 85% chance of meeting its enrollment target if it had worked with a set of 29 sites selected based on database research and different criteria, and if it had allotted 4 months longer for the process.

Future perspective

Recognition of the value of a data-driven approach to performing feasibility analyses is spreading, and the race is on among industry vendors to develop the most comprehensive and accurate feasibility solution. Interest in data-driven feasibility assessment as a business is also being fueled by the growing availability of the data itself; access to structured and unstructured data from public sources is expanding. Competitors in this thriving business will include established companies as well as start-ups.

A key area for exploration of value in this space will be social media. The explosion of online forums for patients and caregivers will make it easier to understand and monitor the intensity of activity around a disease in relation to geography. Also, as a result of more industry consortia and sponsors' interest in sharing data on sites, more information will also become available to study planners about sites' infrastructure and performance. Consequently, sponsors will need to rely less on subjective information supplied by investigators.

Executive summary

Patient enrollment statistics

- The majority of clinical trials deviate from their forecast due to difficulties in enrolling patients in the prescribed time frame.
- The problem is not a lack of planning, but rather of planning based on faulty and/or incomplete information gleaned from investigators about how many patients they can furnish.

Traditional feasibility practices & limitations of relying on feedback from investigators

- Investigators overestimate how many eligible patients they see.
- Surveying investigators, as has been done traditionally, should not be the main source of information on which recruitment plans are based, but one among many.

Conducting due diligence

- With the advent of commercial databases on patient health and investigators' performance, sponsors can conduct a much broader, evidence-based due diligence process to determine if the job can be done on time, where it should be done and by whom.
- The first step is to assess patient availability, by country. Electronic Medical Records, prescription, and integrated medical claims databases can be mined for basic patient counts. These should be adjusted to reflect competition from existing trials in related therapeutic areas. The result is a heat map showing the geographic concentration of target patients.
- Primary research and social media listening should be undertaken to gather the patient perspective and ensure that the protocol does not pose an unforeseen emotional or logistical hurdle for patient participation.
- Countries can be selected based on an analysis of investigator availability, relevant trial experience, site initiation time lines, epidemiology and other patient availability data, competitive trials and standard of care compatibility.
- The heat map can be used, along with data on where sites are located, to identify sites in proximity to the greatest number of patients.
- A list of preferred sites can be further refined by researching their historical performance and capabilities – again via commercially available data sets.
- And lastly, sites can be surveyed as to their capacity to participate in the trial in question. Such surveys can round out the above information gathered from third parties, including sites' own estimates of the number of patients they can provide.

Simulating enrollment rates

- The above quantitative information should then be fed into a statistical model to determine if the right planning parameters are in place. The output of the model is a distribution chart showing the probability of meeting the enrollment target for a given set of variables.
- The trial parameters should be adjusted until the model suggests the level of probability of success that the sponsor can accept.
- This approach has been shown in real-life applications to provide accurate forecasts of what it will take to enroll a trial in a certain time frame, given specific parameters.

Future perspective

- Sponsors that adopt this methodology, leveraging available data and sophisticated analytics stand to benefit from realistic trial forecasts and achievable plans. By improving the predictability of their trial planning, study planners can approach their work with confidence and drastically reduce the amount of bad news they have to deliver.
- This level of precision in enrollment forecasting has never been possible before and stands to improve the trial enrollment process dramatically, positioning clinical trials for success and saving sponsors millions in unexpected costs and delays.

Financial & competing interests disclosure

O Johnson is an employee of inVentiv Health. The author has no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

Papers of special note have been highlighted as:
• of interest

1 CenterWatch.CenterWatch to Discuss critical physician-patient relationship to maximize clinical study referrals at MAGI's Clinical Research Conference. Press release: 2009. www.centerwatch.com/press/pr-2009–10–05.aspx.

• **Tufts and CenterWatch are constantly monitoring the industry, conducting surveys and reporting metrics on performance in this space. We have referenced these sources because there is no better reference source for this type of information.**

2 Kaitin KI. 89% of trials meet enrollment, but timelines slip, half of sites under-enroll. *Tufts Center for the Study of Drug Development Impact Report.* 15(1) (2013). http://csdd.tufts.edu/files/uploads/jan-feb_2013_ir_summary.pdf

• **This study takes a current and thorough look at enrollment time lines, highlighting the need to strategically address ways to accelerate time line to bring medicines to market faster.**

3 Clinical trial delays: America's patient recruitment dilemma (2012). http://www.drugdevelopment-technology.com/features/featureclinical-trial-patient-recruitment/featureclinical-trial-patient-recruitment-1.html

4 Kaitin KI. Fastest drug developers consistently best peers on key performance metrics. *Tufts Center for the Study of Drug Development Impact Report.* 8(5) (2006).

• **This study takes a current and thorough look at enrollment time lines, highlighting the need to strategically address ways to accelerate time line to bring medicines to market faster.**

5 Hess J. Web-based patient recruitment: best opportunity to accelerate clinical trials. Cutting Edge Information. www.cuttingedgeinfo.com/process/?ref=122 .

• **Evaluates the industry's current expenditure on centralized recruitment outreach, as well as performance by tactic and recommendations to ensure success.**

6 The voice of the patient: a series of reports from the FDA's patient-focused drug development initiative. www.fda.gov/forindustry/userfees/prescriptiondruguserfee/ucm368342.htm

7 Citeline. Sitetrove® (2014).

8 United Health Group. Insurance claims database through OptumInsight (2014).