The Use of Biomarkers in Human Pharmacology (Phase I) Studies

A.F. Cohen,1,2 J. Burggraaf,1,3 J.M.A. van Gerven,1,2 M. Moerland,1 and G.J. Groeneveld1,4

1Centre for Human Drug Research, 2333 CL Leiden, The Netherlands; email: ac@chdr.nl
2Leiden University Medical Center, 2333 ZA Leiden, The Netherlands
3Leiden Academic Centre for Drug Research, Leiden University, 2333 CC Leiden, The Netherlands
4VU University Medical Center, 1081 HZ Amsterdam, The Netherlands

Keywords
question-based drug development, surrogate endpoints, clinical trials, precision medicine, prediction in pharmacology

Abstract
The development of a new medicine is a risky and costly undertaking that requires careful planning. This planning is largely applied to the operational aspects of the development and less so to the scientific objectives and methodology. The drugs that will be developed in the future will increasingly affect pathophysiological pathways that have been largely unexplored. Such drug prototypes cannot be immediately introduced in large clinical trials. The effects of the drug on normal physiology, pathophysiology, and eventually the desired clinical effects will need to be evaluated in a structured approach, based on the definition of drug development as providing answers to important questions by appropriate clinical studies. This review describes the selection process for biomarkers that are fit-for-purpose for the stage of drug development in which they are used. This structured and practical approach is widely applicable and particularly useful for the early stages of innovative drug development.
INTRODUCTION

In the 1960s, Black et al. (1) discovered propranolol, the first successful antagonist to the β adrenoreceptor. Techniques for testing the pharmacology of the novel drug in humans were already available (2, 3), and before the drug was tested in patients (4), the effects of the β-blocker on the actions of epinephrine (5) were tested in healthy volunteers after infusions in the brachial artery. Thus, progression of a receptor blocker from molecular pharmacology through animal pharmacological models to human tests using pharmacological challenges to clinical trials in a diseased population is not a new approach.

In the half century following these landmark studies, as drug development progressed and increased in size and complexity, this logical scientific approach was largely abandoned rather than expanded. In an earlier paper, we showed that in the Netherlands, most early studies with new medicines in humans had clinical tolerability as the primary objective rather than an indication of the human pharmacology of the compound (6). The puzzling question that arises is why early studies are not designed to be more informative. The answer may be that most medicines are now developed for chronic diseases including cancer, diabetes, or inflammatory conditions such as atherosclerosis. The discovery of a new drug candidate is based on presumed effects on biochemical or immunological mechanisms, but the important clinical effects can only be discovered over the long term using outcome measures such as improvement in function or survival. Early knowledge of precise, quantitative, pharmacological effects is often not seen as predictive of later clinical events.

An example is the cholesteryl ester transfer protein (CETP) inhibitor torcetrapib (7). Extensive validated evidence that increased high-density lipoprotein (HDL) cholesterol was associated with beneficial effects on survival from atherosclerosis existed when the drug was discovered and developed. Researchers knew that inhibition of CETP would increase HDL (the short-term pharmacological effect), implying that it therefore would reduce mortality in patients with vascular disease (the clinical effect). Initial studies in healthy subjects demonstrated the quantitative relationship between the blood concentration of the CETP inhibitor and CETP inhibition and increases in HDL cholesterol (8). Unfortunately, clinical trials with this compound demonstrated no beneficial clinical effects (7, 9) and had negative effects on cardiovascular events. The development of the drug continued until the definitive clinical studies demonstrated these adverse effects convincingly, and development was stopped at a cost of about $800 million. The reason why an apparently beneficial improvement in a well-known and apparently validated marker of vascular disease (HDL cholesterol) did not lead to clinical improvement has not been fully elucidated but may have been caused by other effects of torcetrapib, notably increases in blood pressure and effects on aldosterone secretion. These effects were potentially detectable in animals (10) and were seen in earlier studies (11) and even in late clinical studies (9, 12).

This therapeutic failure may have added to existing skepticism about the usefulness of collecting early pharmacological data in human studies (13). However, integration of the available data from early human pharmacology could arguably have prevented this mishap. Examples such as this should probably lead to collecting and above all connecting more, rather than less, data on the pharmacological effect profile.

There are several other examples similar to this, such as the CD28 agonist TGN 1412 that caused severe immunological damage (14) and the polymer RheothRx that produced renal damage and excess death (15) in human trials despite earlier available data that indicated this possibility (16, 17). The question of why such obviously negative information obtained in the early stages of research was not adequately investigated will probably never be answered (16) but can be seen as an important motive for a more integrated form of knowledge management in the clinical
INVESTIGATIONAL MEDICINAL PRODUCT

An investigational medicinal product is a pharmaceutical form of an active substance or placebo being tested or used as a reference in a clinical trial. This includes products that already have a marketing authorization but are used or assembled (formulated or packaged) differently from the authorized form, used for an unauthorized indication, or used to gain further information about the authorized form.

development of new medicines. Here we present a comprehensive approach that is applicable to a wide range of new interventions and can be implemented relatively easily. In this review, we use the words drug, medicine, and compound for brevity as synonyms for the term investigational medicinal product used in the regulatory directive (18) (see the sidebar entitled Investigational Medicinal Product).

PHASE I STUDIES AND BEYOND

Traditionally, the process of developing a new medicine is divided into four distinct and consecutive phases, which are described clearly in many textbooks (19). Phase I studies commonly have the simple aim of providing information about tolerability (evaluated by the incidence of subjective side effects), safety (measured by side effects and laboratory values), and the pharmacokinetics of the drug. In many cases, researchers make an attempt to find the maximally tolerated dose. The studies are often done in healthy male subjects, unless inherent toxicity of the compound precludes this, as for the traditional cytotoxic drugs. In Phase II, the diseased subject is studied, and the dose of the drug is titrated against a measure of disease activity. Phase III is for the performance of the controlled trials that lead to registration, and Phase IV trials are those done after marketing approval.

The US Food and Drug Administration (FDA) and the European Medicines Agency (EMA) issued a guideline on general aspects of clinical trials in 1998 that made the clear point that a linear approach to drug development is not recommended (20). The regulatory guidance starts with the statement that “The essence of rational drug development is to ask important questions and answer them with appropriate studies” and subsequently gives the rationale for using a system of classification based on objectives rather than the passage of time (20). The regulatory authorities describe drug development as an adaptable series of studies in which, as in all scientific evaluation, previous experiments determine the nature of the subsequent experiments. Such an approach is incompatible with the phased approach, in essence a linear time series of experiments. In fact, a Phase III trial may easily generate further preclinical or even (traditional) Phase I trials; the CETP inhibitor torcetrapib, mentioned above, is a case in point, as the findings in late-stage trials prompted further preclinical work to elucidate the mechanism of the off-target effects on blood pressure. The choice between a phased approach and a more cyclical form of project management in innovation has been the subject of much work in the management literature (21, 22). It is generally accepted that a phased approach is appropriate for the standard development of incrementally innovative projects, in which the expected end result is known and few unknown findings are expected. In the 1998 guideline, the EMA had already stipulated that new drug development required a more cyclical approach than the hitherto-accepted phased one (20).

This guideline is generally ignored, even by the regulatory agencies themselves, who have adhered to classifying research according to the passage of time rather than objectives and have added more subphases (0, Ia, and IIb). We have argued earlier that, in early phases of development, this classification is uninformative and that the objectives of tolerability and safety in human
pharmacology studies cannot be attained because of the small number of subjects exposed in these early trials, unless unexpected severe toxicity occurs (23). In this review, we use the preferred terms from the EMA guideline, based on the objective of the studies (human pharmacology, therapeutic exploratory, and use studies), rather than Phases I–IV. Our focus is the establishment of a logical and practical system for the selection and use of biomarkers in human pharmacology and therapeutic exploratory studies.

THE DRUG PROTOTYPE

The rational development of new medicines, as an adaptable series of studies designed to answer important questions (20), is most applicable to a subset of innovative products with a novel mechanism of action. In these situations, the uncertainty about the different aspects of the drug is the greatest, and flexibility in the drug development program and innovative methods are the most urgently required.

The definition of this subset of projects is important for research organizations because the project management is radically different depending on the type of compound. The innovative nature of a compound depends on two continuous variables (6). One is the level of certainty about the link between the pharmacological/physiological mechanism and the intended clinical effect. The other is the level of validation of a measurement or set of measurements that indicates the strength of this link. When both are at a high level, as in the case of a generic version of an existing effective drug from a well-established class, there are a limited number of unknown areas that can be managed easily during development. Such compounds can be developed in a highly organized, time-phased fashion and are not the subject of this review. The project management of innovative, high-uncertainty projects requires different management, staff competences, and systems that are largely related to the integration of information rather than the purely operational requirements of a generic development program (22).

OVERVIEW OF DEVELOPMENT MODELS

Many models for drug development have been described (see Figure 1). By the time the FDA guideline was released, Sheiner (24) had already introduced a cyclical model of development that recommended cycles of learning and confirming research. He emphasized that much can be learned from existing data by applying mathematical models that still assume a time-phased approach with all measurement techniques to attain the goals already in place. In a paper on mechanism-based modeling, Danhof et al. (25) proposed seven classes of measurements that could be used in clinical development and recommended that these be used in mathematical models to assist in decision making in the development process. This classification is conceptually interesting but has the drawback that the classes are not mutually exclusive and are represented in a linear manner, as if one would always follow the other.

Lalonde et al. (26) and Milligan et al. (27) described an integrated, model-based drug development paradigm in which all aspects of the development program, from modulation of a biological

Figure 1
Models of development and biomarker classification. (a) The learn-and-confirm paradigm from Sheiner, adapted from Reference 26 with permission. (b) The mechanism-based biomarker classification by Danhof et al. (25). (c) Model-based drug development by Lalonde et al. (26) and Milligan et al. (27). Abbreviations: M&S, modeling and simulation; PD, pharmacodynamics; PK, pharmacokinetics. Panel b adapted from Reference 25 with permission. Panel c adapted from References 26 and 27 with permission.
a. Drug development and model building

Continuum of learn/confirm/predict at each decision point

M&S

Preclinical Phase I Phase IIa Phase IIb Phase III Registration

Efficacy

Toxicology

PK/PD

Toleration

Human PK/PD

Efficacy and safety

Dose or exposure-response

Therapeutic index

Covariate effects

Results relative to competitors,

regional differences,

and therapeutic index

Uncertainty

Confidence in drug and disease

b.

Type 0

Phenotype/genotype

Type 1

Drug concentration

Type 2

Target occupancy

Type 3

Target activation

Type 4

Physiological response

Type 5

Pathophysiological response

Type 6

Clinical response

Rationale #1

Pathway

Right pathway

Systems biology

Impact

Rationale #2

Target

Right target

Systems pharmacology

Rationale #3

Drug

Right molecule

Translational sciences

Rationale #4

Benefit/risk

Right dose

Exposure response

Rationale #5

Effectiveness and reimbursement

Optimized medicines

PK/PD

Concentration

Time

Effect

PD

www.annualreviews.org • Use of Biomarkers in Human Pharmacology 59
pathway to financial reimbursement on the basis of cost effectiveness, can be included. They demonstrated that a quantitative approach, using mathematical models that allow simulation of research steps, can lead to increased efficiency.

All these models are conceptually elegant and statistically and mathematically complex, but they depend fully on what is measured and how well such measurements reflect the objectives of the research or, in the words of the FDA guideline, how well they address certain questions. Therefore, applying the inappropriate measurement to a particular question can result in an unsuccessful development trajectory.

In the 1990s, a generally accepted measure for cardiac arrhythmia—the number of ventricular extrasystoles during continuous ECG recording—was reduced effectively by several drugs, but all increased mortality (28). This was an excellent demonstration of how a pharmacological mechanism (interference with Na\(^+\) channels in the cardiomyocyte) affected a measurement of organ function (number of extrasystoles) but did not correlate with the clinical endpoint (improvement of survival). Although it has often been stated that this result invalidates the measurement of extrasystoles, this is only true in the sense that they are not a valid surrogate for the effect of the drug on survival. However, they are a very good measure for the drug’s effect on Na\(^+\) channels. In the majority of subjects, there is no direct association between extrasystoles and survival. Another example of pharmacological effects being used erroneously as a surrogate was the dopaminergic drug ibopamine. Introduced in the 1980s with positive effects on renal and cardiac hemodynamics in heart failure (29, 30), it was marketed for a long time between 1991 and 1995 before trials of mortality and morbidity first raised doubts (31). Subsequent trials with mortality as the endpoint demonstrated that the drug was actually deleterious (32), and ibopamine was withdrawn from the market. Clearly, the short-term improvement of heart function was not a surrogate for improvement of the mortality rate (33). This demonstrates the importance of using a model that indicates what questions have to be asked, what methodology is used for providing the answers, and how the progression to establishing the efficacy and safety of a medicine is mapped and communicated.

A NEW, PRACTICAL, CONCEPTUAL MODEL FOR EARLY PROTOTYPE DEVELOPMENT

Any proposed conceptual approach to the development of new medicines is a method of communication within a project team, from the project team to management and funders and to authorities who judge the value of a new treatment. Moreover, there is an increasing need for communication to the patients or their representatives, who will rightly want to play a role in the development. Any model should therefore be relatively simple and provide continuous opportunities for tracking progress.

Our proposed approach starts with the determination of the prototypical nature of the drug candidate (Figure 2). In practice, this means that the validation of the target must be judged against the validation of the available measurement systems. This model was introduced as a conceptual model of development (6), indicating the need for validation of the targets as well as the biomarkers. This model is supplemented by the model of Danhof et al. (25), which describes a logical nomenclature of biomarker types linked to their purpose for biomarkers of drug activity, and the model-based drug development approach described by Milligan et al. (27) and Lalonde et al. (26). The synthesis of models we now describe may be a useful tool to plan development and track progress. Moreover, this synthesis makes clear that drug development is not a direct trajectory from a prototype to a drug in practical use. We have reported earlier that in a set of human pharmacology (Phase I) studies submitted to the Netherlands Clinical Trial Competent...
Figure 2
Identification of prototypical projects. This model incorporates the drug development models of Danhof et al. (25), Milligan et al. (27), and Cohen (6). The thick white line indicates a perceived ideal drug development trajectory in which, after identification of a suitable drug molecule, trials are performed that show clinical utility. Unfortunately, the situation is often more like that of the antiarrhythmics in the Cardiac Arrhythmia Suppression Trial (CAST, red circle) (28), in which a target was chosen (Na+ channels in the myocyte) and an inhibitor (flecainide) and a physiological biomarker (extrasystole on the ECG) were established. The final trial incorrectly assumed that this was a clinical biomarker, and the trajectory (thin white line) turned toward a late-stage failure (LSF, yellow circle), causing doubts about whether the pathway and target were correct for the clinical endpoint (mortality). The blue line indicates a hypothetical program with learn-and-confirm cycles at many points, biomarkers that are tuned to the objective of the development stage, and a longer, but in this case successful, development program with reduced risks. Note that in the earlier phases, a loop (dashed line) back to the prototypical state—in that case, often for a different indication—is possible. Figure adapted from References 6, 25, and 27 with permission.

Authority, the use of early quantitative markers of drug action was still very low in 2009 (6). Many examples from the past, as presented in this review, do not support this fast-lane approach for prototypes. The model presented in Figure 2 allows investigators to make the prototypical state of a new drug explicit. If the link between the biological mechanism and a disease is already strongly established and good biomarkers are available for all stages, development can be standard. In other cases, the nonstandard approach must be determined on a case-by-case basis. The next section presents a practical approach to this.

DEFINITIONS OF BIOMARKERS
Beginning around 1990, drug development scientists were increasingly interested in quantitative markers of disease activity, perhaps influenced by the need to obtain rapid quantitative indications...
of drug action (34) to speed the development of new anti-HIV agents. The importance of these quantitative markers increased rapidly. The number of publications with biomarker as a search term went up about 6-fold between 1985 and 2010, while the total number of publications only doubled.

The most useful definitions of the terms currently available are the result of a report by the National Institutes of Health Biomarkers Definitions Working Group in 2001 (35):

- **Biological marker (biomarker):** a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacological responses to a therapeutic intervention;
- **Clinical endpoint:** a characteristic or variable that reflects how a patient feels, functions, or survives; and
- **Surrogate endpoint:** a biomarker that is intended to substitute for a clinical endpoint. A surrogate endpoint is expected to predict clinical benefit (or harm or lack of benefit) based on epidemiological, therapeutic, pathophysiological, or other scientific evidence.

These definitions have provided considerable clarity to the field. A biomarker is just a quantitative indicator of a biological process. In drug development and pharmacology, influencing such a process is done only with the goal of attaining a clinical endpoint, which is an improvement of the functioning, feelings, or survival of a patient. Clinical measurement of these effects tends to be variable and is often related to prevention of delayed but serious events, such as mortality, or disease manifestations, such as stroke or infarction. This means that meaningful effects on such measurements require very large and/or long trials (36). Therefore, a parametric and accurate marker of the pharmacological influence of a critical pathophysiological process should be able to reflect the clinical effect to such an extent that this measurement could replace the clinical outcome measurement. Such a measure is then called a surrogate endpoint. This ultimate, elevated state of a biomarker is reached rarely. Perhaps the best examples are the serum viral load in HIV or hepatitis C. The state of surrogacy for a biomarker can only be assured when the clinical program for a drug has been completed, and the earlier surrogate biomarker is only useful for the faster development of similar medicines.

The need for more advanced measurement tool kits in clinical research was highlighted in a 2004 report by the FDA (37), which issued subsequent guidance on the validation of biomarkers in both genomics and biochemistry (38, 39). Lee et al. (40) provided useful clarifications about validation, which is usually defined as assessing the performance characteristics of a biochemical assay. For a biomarker to have utility in drug development and for acceptance for registration purposes, the traditional requirements for sensitivity, specificity, and precision are just as important as for any laboratory test. The clinical validation of biomarkers as surrogate markers involves a process for which the FDA introduced biomarker qualification as the preferred term (35).

**BIOMARKERS PROVIDE THE ANSWERS TO SCIENTIFIC QUESTIONS**

The definition above does not specify what question a particular measurement is intended to address. The catalogue of potential measurements in biology is virtually unlimited; therefore, the choice of biomarkers requires guidance. The basis for this can be found in the EMA guideline about clinical trials (20), which states that rational drug development asks important questions and answers them with appropriate studies. We now expand this definition as follows:

*The essence of rational drug development is to ask important questions and answer them with appropriate studies using fit-for-purpose, qualified biomarkers.*
QUESTION-BASED DRUG DEVELOPMENT

The next step in every drug development program is to determine what these important questions are. The objectives from the guidance are presented in a manner that appears generally applicable (see Table 1), and it is therefore attractive to use this approach for a drug with a known mechanism of action and a development trajectory that has been explored previously. Such nonprototypical development is relatively low risk.

However, for prototypical compounds with new proposed mechanisms, standard development may lead to serious oversight of essential problems. The most well-known example is TGN 1412, a drug with an entirely new immunological mechanism, for which questions about the pharmacology of the drug in humans were obvious before the trial. Such questions could have been answered by simple biomarkers obtainable ex vivo and by calculations of receptor density that would have revealed that this drug was overdosed 100–1,000-fold during the first administration to humans and consequently caused serious harm to the subjects (14, 41).

Another example is the 5-hydroxytryptamine4 (5-HT4) receptor agonist tegaserod, which was introduced for the treatment of irritable bowel syndrome in 2001 and was taken off the market in 2007, when a meta-analysis of adverse events showed a small but significant increase in cardiovascular events. The drug was found to be well tolerated in healthy volunteers (42), and the compound underwent extensive research that was all directed toward the effects on the colon. When the drug was withdrawn, several small studies were published (43–45) that raised doubt about the presence of any cardiovascular effects of the drug. However, investigators knew that tegaserod was pharmacologically nonselective and its effects on 5-HT4 receptors on platelets may have contributed to the cardiovascular events (46). It is tempting to speculate what would have happened if the question about 5-HT4 antagonism had been asked at the start of development and if adequate studies to address the cardiovascular risks had been carried out properly.

Table 1  An approach to classifying clinical studies according to objective

<table>
<thead>
<tr>
<th>Type of study</th>
<th>Objective of study</th>
<th>Study examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human therapeutic</td>
<td>Assess tolerance; define/describe PK and PD</td>
<td>Dose-tolerance studies</td>
</tr>
<tr>
<td>Pharmacology</td>
<td>Explore drug metabolism and interactions</td>
<td>Single- and multiple-dose PK and PD studies</td>
</tr>
<tr>
<td></td>
<td>Estimate activity</td>
<td>Drug interaction studies</td>
</tr>
<tr>
<td></td>
<td>Explore use for targeted indication</td>
<td>Earliest trials of relatively short duration in well-defined, narrow patient populations, using surrogate or pharmacological endpoints or clinical measures</td>
</tr>
<tr>
<td></td>
<td>Estimate dosage for subsequent studies</td>
<td>Dose-response exploration studies</td>
</tr>
<tr>
<td></td>
<td>Provide basis for confirmatory study design, endpoints, and methodologies</td>
<td></td>
</tr>
<tr>
<td>Therapeutic</td>
<td>Demonstrate and confirm efficacy</td>
<td>Adequate, well-controlled studies to establish efficacy</td>
</tr>
<tr>
<td>Exploratory</td>
<td>Establish safety profile</td>
<td>Randomized parallel dose–response studies</td>
</tr>
<tr>
<td></td>
<td>Provide an adequate basis for assessing the benefit-risk relationship to support licensing</td>
<td>Clinical safety studies</td>
</tr>
<tr>
<td></td>
<td>Establish dose-response relationship</td>
<td>Studies of mortality/morbidity outcomes</td>
</tr>
<tr>
<td></td>
<td>Refine understanding of benefit-risk relationship in general or special populations and/or environments</td>
<td>Large, simple trials</td>
</tr>
<tr>
<td></td>
<td>Identify less common adverse reactions</td>
<td>Comparative studies</td>
</tr>
<tr>
<td></td>
<td>Refine dosing recommendation</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: PD, pharmacodynamic; PK, pharmacokinetic.

aData from Reference 20.
The question-based drug development approach shown in Figure 3 has been described in detail by de Visser (47). When a new project is initiated, this approach allows a project leader or team to develop an initial outline of the important questions. The figure shows a hypothetical example for a central nervous system (CNS) drug, but the questions have to be adapted to the actual therapeutic situation. The questions are arranged in six domains that follow the logical progression of a drug through the body, from absorption to excretion and the pharmacological, (patho)physiological, and eventually clinical effects that may occur. In an important study, researchers at a major pharmaceutical company found that in 43% of projects, the mechanistic aspects of new compounds were not tested adequately. They called knowledge about (a) exposure at the site of action, (b) target binding, and (c) the establishment of functional pharmacological activity the three pillars of survival of a new compound (48). Drug development without this knowledge had a considerable chance of failure. These pillars were built into the question-based scheme. This model can be used for individual studies, for programs and projects, or both.
From Questions to Answers

After the questions have been defined, generally in a process that involves several meetings, the next step is to determine which biomarkers are available for obtaining the answers. Additionally, the qualification and validation status of the biomarkers has to be assessed. This is important because many measurements are unsuitable for use in a particular type of trial, even though they are adequately described in the literature. Often, a considerable amount of work is required to make them fit-for-purpose (answering the important question).

Conversion of Questions to Studies and Real Options Project Planning

The set of questions defined above now needs to be translated into a series of studies or a single study. The advantage of this approach is that the objectives of the studies can be derived easily from the questions. The order in which the questions are answered is often apparent but can also be determined more formally by decision analysis based on the real options method. This form of decision analysis optimizes the balance between cost and reduction of uncertainty and provides a rational method for the design of a development program (47, 49, 50). The details of this approach are beyond the scope of this review; briefly, they require a priori assessment of the cost of answering a certain question and the probability that useful information about this question will be obtained. A real options decision analysis then provides a logical order that optimizes information content and cost. Quantitative methods for project planning have been underused in pharmaceutical research. The main advantage of the real options method is that it makes the value of scientific information explicit rather than just expressing it as a cost.

The development trajectory has now been turned into a question-based plan supported by a series of studies and the possibility of tracking the progress through real milestones rather than anonymous phases in time. All that remains is for researchers to choose the right biomarkers.

SAFETY BIOMARKERS

We do not make a special case for biomarkers of adverse effects or safety. Aronson & Ferner (51) have proposed the term off-target pharmacology, recognizing that adverse events are also reflections of the pharmacology of the compound. This is even true for rare but serious events such as anaphylaxis or severe liver damage induced by drugs, although sometimes the mechanism is not known. The development of a prototypical compound is always based on some balance between target and off-target effects, and the question-based scheme makes that explicit. Especially in early development, the same biomarkers can reflect, in some cases, an off-target effect and, in other cases, the desired pharmacology of a drug. An example is the centrally acting antihypertensive rilmenidine, for which the speed of horizontal eye saccades was used to demonstrate off-target sedative effects on the CNS (52). For the orexin antagonist almorexant (53), the same biomarker was used to quantify the intended (on-target) pharmacological effect on sleepiness.

Another aspect of the fit-for-purpose approach to biomarkers for off-target effects is that frequent measurement can enhance their usefulness. Supplemental Figure 1 shows the effect of a new antisense compound (54) on renal biomarkers as an example (follow the Supplemental Material link from the Annual Reviews home page at http://www.annualreviews.org).

QUALIFICATION OF BIOMARKERS BY PHARMACOLOGY

In its early phase, drug development research is conceptually no different than late-phase research. However, in the early phases of research with prototypical drugs, the biomarker plan will include a
much greater proportion of biomarkers that are not yet qualified, as the need for measurements of certain variables only arises because a prototypical drug affects a hitherto-unaffected mechanism. The measurements are therefore just as unknown as the new mechanism and their relevance still uncertain. This is probably one of the reasons investigators are reluctant to use biomarkers in early-phase studies. However, obtaining the most detailed information about the pharmacology of a new drug in humans is obviously also of the utmost importance. For the less innovative compounds, a course of action can readily be found in the relevant literature and regulatory documents and, if necessary, adapted. The incorporation of biomarkers in early development of prototypes sometimes requires the development of unqualified biomarkers to quantify an unknown mechanism, and this is not possible. In principle, the biomarker must be established as useful before the drug can be evaluated in its different aspects. This is clear from Figure 2, in which the trajectory proceeds upward (develop the biomarker) before moving to the right (develop the drug).

Mechanism-Specific Biomarkers

Mechanism-specific biomarkers are also sometimes called proximal (55) and reflect the action of a drug on pharmacology or pathophysiology (types 0–3 in Figure 2). The requirements for validation and qualification are well established and described for biomarkers that are collected from biological fluids and measured in a biochemical or genetic laboratory (sometimes called wet biomarkers) (40, 55). Most of the criteria can also be applied to biomarkers obtained from imaging with positron emission tomography, which measures in vivo binding of a drug to a particular receptor. The requirements for such measurements are also relatively easily derived from the mechanism of action and the off-target effects of the new medicine. This can be done on the basis of the question-based plan. Additionally, these biomarkers can often be tested after animal experiments, and experimental drugs can be added to human blood and examined in the laboratory before any trials in human subjects begin. Perhaps the most widely used biomarker is the concentration of the drug and its metabolites in biological fluids, for which the availability of a reliable and rapid assay during early stages of development in human trials is essential. In the traditional approach, this would be just a marker for pharmacokinetics. However, as a measure that is obtained after each dosing step, it provides an immediate link to the toxicological and pharmacological experiments in animals and can be used for mathematical modeling of the drug-target interaction (27, 56).

An example of this is the development of a low-molecular-weight inhibitor of mitogen-activated protein kinase (Supplemental Figure 2). Such an approach could be a typical one for question-based development with built-in learn-and-confirm cycles. Any biomarker qualification requires time, and this adds strong support for a question-based biomarker plan that is carried out as early as possible in development and well before first-in-human studies.

Clinical Biomarkers

Validation and qualification are less obvious for clinical biomarkers than for biomarkers demonstrating the pharmacology of a drug. Clinical biomarkers (also termed distal; types 4–6 in Figure 2) are thought to reflect disease activity and pathophysiology (55). Some clinical biomarkers are clearly validated, but many of these relate to the delayed occurrence of events, such as mortality or the incidence of certain adverse events.

Selection of clinical biomarkers. Although it is possible that, for some effects, a whole new test must be constructed, functional tests of most organs and physiological systems are generally
available. Rather, there is often an embarrassment of choices, and their usefulness has generally been determined in clinical practice, where they are used as markers for disease progression. For example, plasma creatinine concentration is a widely used marker for disease progression in renal disease, but it lacks utility for the detection of potential, small changes in renal function induced by new medicines (54, 57). Simply collecting such functional biomarkers from the literature does not provide a useful basis for selection because there is no indication of the performance of such a biomarker in drug studies. We have published several meta-analyses of biomarkers based on the assumption that, when a biomarker is sensitive to the effects of compounds from a certain mechanistic class, this will also hold for subsequent new compounds from that class. These analyses were mainly done in the CNS area for benzodiazepines (58), antipsychotics (59), cannabinoid drugs (60), selective serotonin reuptake inhibitors (61), and alcohol (62). In the CNS area, the choice of potential biomarkers of CNS function is virtually unlimited, and hundreds of tests and test variants have been published. The selection process used in this area can also be transferred to other areas, such as cardiovascular or pulmonary functional testing, where the number of tests is generally smaller and the problem easier to handle.

Selection process. A useful biomarker in any class has to comply with the following general criteria (58):

- There must be a consistent response of the biomarker across studies (preferably from different research groups) and drugs from the same mechanistic class.
- The biomarker must respond clearly to therapeutic (not supratherapeutic) doses.
- There must be a clear dose- or concentration-response relationship.
- There must a plausible relationship between the biomarker, pharmacology of the drug class, and disease pathophysiology.

An example of how the selection process takes place can be found in several meta-analyses specifically performed for this purpose (58, 60, 61, 63).

Validation. The meta-analysis technique is very useful for the selection of biomarkers that already exist and can be used to highlight a certain type of test. However, this does not mean that an identified test is immediately usable. Even when the four criteria indicated above apply, qualification of a test in real-life circumstances is indicated. For a laboratory-based biomarker, a well-described validation plan is available using quality control samples, calibration, and tests of accuracy and precision (40). In principle, such tests could be performed with a biomarker measuring clinical function, but in practice, this is not practical; each step requires a clinical trial, and multiple positive and negative controls would have to be tested. Despite this, at least some steps have to be taken, including determining the sensitivity of a test.

The sensitivity of a test can vary considerably. For instance, the well-known sedative benzodiazepine temazepam produced significant negative results on a particular eye-hand coordination test, the pursuit rotor, only after a dose of 40 mg (64), whereas the effects of a dose of 5 mg were easily detectable by a reduction of the speed of saccadic eye movements (65). If temazepam had been an experimental drug and the insensitive test was used during early development, the active dose would have been grossly overestimated. Although the most sensitive tests will generally reveal more consistent significant effects in a literature overview or meta-analysis, sensitivity studies with relevant comparator drugs can greatly increase the confidence in a biomarker, even though such studies are time-consuming. The selectivity of a test can be determined by using it on a range of drugs affecting a certain function but acting through different mechanisms. This is particularly relevant for CNS medicines, where the functions to be measured are fairly generic and related to
activation or depression of brain function, but the mechanisms from which these effects originate vary considerably. A practical approach is to evaluate several known drugs and perform a range of tests to see which ones respond. For instance, this was done with a new magnetic resonance imaging measure (resting-state network analysis) by comparing alcohol and morphine—both have a known sedative potential but also act through diverse mechanisms (66). Such a clinical biomarker validation program would reveal if tests that were described as useful do not in fact meet the criteria for a biomarker (67).

**Practical validation.** Investigators must also pay attention to the practical aspects of a clinical biomarker. A test usually has to be administered multiple times to follow the effect of a drug and must not be too long or too burdensome for the subjects, who might lose motivation to participate. Evaluation of the test results generally must be fast, and automated and secure data collection is therefore important. The ability to study several subjects at the same time requires test instruments that are well characterized and technically standardized so that multiplication of instruments does not become unduly expensive and, when necessary, the test system can be transferred to other study sites. In addition, attention must be given to the reliability of a biomarker in clinical trials in which the circumstances and patient populations may require adaptations of the tests.

**Use for new medicines with an untested mechanism of action.** The indicated approach may lead to the best possible clinical biomarker set. However, for a medicine with an entirely novel mechanism of action, it is not certain that these biomarkers will actually give the desired response.

The biomarker could be unresponsive to the effect of the drug because another pathway or function is affected. For instance, cannabinoid agonists do not affect the speed of saccadic eye movements at all, but they have a pronounced effect on subjective alertness and body stability (68). Selective γ-aminobutyric acid (GABA) agonists have little effect on alertness and body stability, whereas the decrease in the speed of saccadic eye movements is similar to that of the nonselective GABA agonist lorazepam (69).

Such differences cannot be immediately predicted unless a drug with a similar mechanism of action has already been tested. Evidence has now been gathered that firmly relates peak saccadic velocity to the pharmacological activity of partial subtype-selective GABA\_A agonists (70). For first-in-class drugs, the only practical solution is to apply a range of tests broadly encompassing the functional domains. Preferably, such tests should evaluate drugs with as much overlap with underlying mechanisms as possible.

Another possibility is that a compound does not produce any functional effects by itself. This is most likely the case for agonists or antagonists of a phasic rather than tonic physiological system or systems with low intrinsic activity. These can only be detected by challenging the system with an agonist. Examples are scopolamine for the cholinergic system in the brain (71), tetrahydrocannabinol for testing cannabinoid antagonists (72, 73), vasopressin for antagonists of the vasopressinergic system (74, 75), and 5-HT for the serotonergic system in the brain (76). In addition, serotonin-induced arterial constriction (77) and the use of 5-HT (78) for the study of drug effects on platelet reactivity are good examples of the use of challenges in vascular pharmacology.

The development and selection of a challenge agent have to comply with the same rules as a direct functional biomarker. The development of such tests for a new compound generally requires time and an early biomarker plan.

**In-trial pharmacological qualification.** The measures described above help to develop a biomarker that is maximally useful and well qualified for the purpose that is predefined in the question-based plan. Yet much depends on the decisions taken in early development, and these
need to be supported by robust data. Biomarkers of function may lose reliability in a trial for many technical reasons or through changes in motivation of subjects or staff. In any early trial, a positive control group treated with a relevant drug of which the performance of the biomarker is known is essential and provides considerable comfort about the findings obtained for a new drug (53, 79). Positive controls can be easily added, even to a first-in-human design, either as a crossover treatment or as a separate group. The positive control group also provides immediate relevance to the size of the functional biomarker’s effect.

Pharmacokinetic-pharmacodynamic qualification. When a measurement with a functional or mechanistic biomarker can be related closely to the plasma concentrations of the drug through a plausible pharmacokinetic-pharmacodynamic model, this implies that the marker is relevant for evaluation of target engagement of the drug. Researchers must still make explicit for which purpose the use of the biomarker is intended. For instance, for the experimental CETP inhibitor BAY 60–5521, investigators elegantly showed that inhibition of the target protein was related to plasma concentration in a pharmacological model. They could also predict how this would affect HDL cholesterol (80, 81). The biomarkers indicate that the drug affects a physiological response but obviously do not provide direct information about the clinical effect. In view of the known problems with blood pressure increases caused by these drugs, it is interesting that blood pressure was used as a safety biomarker and not reported in relation to the plasma concentrations (as an off-target functional marker).

Use of qualified biomarkers in confirmatory clinical trials. The use of biomarkers for target engagement in early-phase trials has important implications for the later-phase drug development. When used in these later-phase trials, researchers can show that target engagement has also occurred in the relevant patient population under more realistic conditions. Biomarker use can lead to enrichment of a patient population by using target engagement as a covariate in the cases in which this is variable for pharmacokinetic or pharmacodynamic reasons. Finally, the use of biomarkers in these trials will contribute to the development of early biomarkers with relevance for clinical effects.

CONCLUSIONS

A summary of the steps that have to be taken before a biomarker can be used in early drug development is given in Figure 4. Quantitative evaluation of the dose-response relationship is the basis of pharmacology and therefore of the scientific development of new investigational medical products in humans. Researchers can quantify the drug response with a multitude of measurements that have to be selected, evaluated, validated, and qualified before a new drug development program is started. Lack of attention to this methodological support for a development program can be expensive and dangerous. We propose a rational and transparent approach to the involvement of important biomarkers in the development of new medicines. Biomarkers should be fit-for-purpose to show essential drug properties and effects. The subsequent application of qualified biomarkers in confirmatory clinical trials may demonstrate whether the underlying pharmacological effects are indeed maintained throughout the development. The FDA succinctly stated the essence of rational drug development as asking important questions and answering them with appropriate studies (20). In this review, we have provided a conceptual approach to populate these studies with appropriate biomarkers.
Overview of biomarker development. The development of a question-based plan and a provisional biomarker plan with a known validation status allows investigators to subdivide the plan into mechanistic biomarkers (generally requiring traditional laboratory validation) and clinical biomarkers (requiring fit-to-purpose qualification). A meta-analysis of available biomarkers needs to be done to seek available knowledge about sensitivity and specificity. On the basis of this analysis, researchers can design further studies to attain a question-based and fit-to-purpose biomarker set. This analysis will also reveal the need for full biomarker development, with implications for the timing of development.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review. All authors are full-time employees of the Foundation Centre for Human Drug Research. The terms of employment of this organization prohibit shareholdings in health-care-related companies. Fees for lectures and other compensation are charged through the Foundation and not directly reflected in the compensation of the employees.

LITERATURE CITED


Contents

A Conversation with Susan Band Horwitz
Susan Band Horwitz and I. David Goldman ........................................ 1

Introduction to the Theme “Precision Medicine and Prediction in Pharmacology”
Paul A. Insel, Susan G. Amara, and Terrence F. Blaschke ......................... 11

Identifying Predictive Features in Drug Response Using Machine Learning: Opportunities and Challenges
Mathukumalli Vidyasagar ................................................................. 15

Predicting Toxicities of Reactive Metabolite–Positive Drug Candidates
Amit S. Kalgutkar and Deepak Dalvie ............................................. 35

The Use of Biomarkers in Human Pharmacology (Phase I) Studies
A.F. Cohen, J. Burggraaf, J.M.A. van Gerven, M. Moerland, and G.J. Groeneveld ................................................................. 55

Improving Postapproval Drug Safety Surveillance: Getting Better Information Sooner
Sean Hennessy and Brian L. Strom .................................................. 75

Preemptive Clinical Pharmacogenetics Implementation: Current Programs in Five US Medical Centers
Henry M. Dunnenberger, Kristine R. Crews, James M. Hoffman, Kelly E. Caudle, Ulrich Broeckel, Scott C. Howard, Robert J. Hunkler, Teri E. Klein, William E. Evans, and Mary V. Relling .................................................. 89

A Personalized Medicine Approach for Asian Americans with the Aldehyde Dehydrogenase 2*2 Variant
Eric R. Gross, Vanessa O. Zambelli, Bryce A. Small, Julio C.B. Ferreira, Che-Hong Chen, and Daria Mobly-Rosen .................................. 107

Pharmacogenomics in Asthma Therapy: Where Are We and Where Do We Go?
Heung-Woo Park, Kelan G. Tantisira, and Scott T. Weiss .............................. 129
Drug Disposition in Obesity: Toward Evidence-Based Dosing

Catherijne A.J. Knibbe, Margreke J.E. Brill, Anne van Rongen, Jeroen Diepstraten, Piet Hein van der Graaf, and Meindert Danhof ......................................................... 149

How Good Is “Evidence” from Clinical Studies of Drug Effects and Why Might Such Evidence Fail in the Prediction of the Clinical Utility of Drugs?

Huseyin Naci and John P.A. Ioannidis ................................................. 169

The Ethics and Economics of Pharmaceutical Pricing

Sara Parker-Lue, Michael Santoro, and Greg Koski ........................................ 191

Brown, Beige, and White: The New Color Code of Fat and Its Pharmacological Implications

Alexander Pfeifer and Linda S. Hoffman .................................................. 207

Candidate Drug Targets for Prevention or Modification of Epilepsy

Nicholas H. Varvel, Jianxiong Jiang, and Raymond Dingledine .................. 229

The Design of Covalent Allosteric Drugs

Ruth Nussinov and Chung-Jung Tsai .................................................... 249

Therapeutic Modulation of Urinary Bladder Function: Multiple Targets at Multiple Levels

Martin C. Michel ....................................................................................... 269

Mineralocorticoids in the Heart and Vasculature: New Insights for Old Hormones

Achim Lother, Martin Moser, Christoph Bode, Ross D. Feldman, and Lutz Hein .... 289

Unraveling the Physiological Complexities of Antibiotic Lethality

Daniel J. Dwyer, James J. Collins, and Graham C. Walker ......................... 313

Intraclass Differences Among Antihypertensive Drugs

R.D. Feldman, Y. Hussain, L.M. Kuyper, F.A. McAlister, R.S. Padwal, and S.W. Tobe ............................................................... 333

Targeting Hsp90/Hsp70-Based Protein Quality Control for Treatment of Adult Onset Neurodegenerative Diseases

William B. Pratt, Jason E. Gestwicki, Yoichi Osawa, and Andrew P. Lieberman .... 353

New Approaches to Inhibiting Platelets and Coagulation

Jeremiab P. Depta and Deepak L. Bhatt .................................................. 373

DREADDs (Designer Receptors Exclusively Activated by Designer Drugs): Chemogenetic Tools with Therapeutic Utility

Daniel J. Urban and Bryan L. Roth ....................................................... 399

Learning by Failing: Ideas and Concepts to Tackle γ-Secretases in Alzheimer’s Disease and Beyond

Bart De Strooper and Lucia Chávez Gutiérrez .......................................... 419
Therapeutic Applications of Extracellular Vesicles: Clinical Promise and Open Questions
Bence György, Michelle E. Hung, Xandra O. Breakfield, and Joshua N. Leonard 439

Eph Receptors and Ephrins: Therapeutic Opportunities
Antonio Barquilla and Elena B. Pasquale 465

Designed Ankyrin Repeat Proteins (DARPs): Binding Proteins for Research, Diagnostics, and Therapy
Andreas Plückthun 489

Synthetic Lethal Vulnerabilities of Cancer
Ferran Fece de la Cruz, Bianca V. Gapp, and Sebastian M.B. Nijman 513

Calcitonin Gene-Related Peptide (CGRP): A New Target for Migraine
Andrew F. Russo 533

Activation and Regulation of Caspase-6 and Its Role in Neurodegenerative Diseases
Xiao-Jun Wang, Qin Cao, Yan Zhang, and Xiao-Dong Su 553

Constellation Pharmacology: A New Paradigm for Drug Discovery
Russell W. Teichert, Eric W. Schmidt, and Baldomero M. Olivera 573

DNA Methylation and Its Implications and Accessibility for Neuropsychiatric Therapeutics
Jeremy J. Day, Andrew J. Kennedy, and J. David Sweatt 591

Targeting Receptor-Mediated Transport for Delivery of Biologics Across the Blood-Brain Barrier
Jason M. Lajoie and Eric V. Shusta 613

Novel Targeted Therapies for Eosinophil-Associated Diseases and Allergy
Susanne Radonjic-Hoesli, Peter Valent, Amy D. Klion, Michael E. Wechsler, and Hans-Uwe Simon 633

Indexes
Cumulative Index of Contributing Authors, Volumes 51–55 657
Cumulative Index of Article Titles, Volumes 51–55 661

Errata
An online log of corrections to Annual Review of Pharmacology and Toxicology articles may be found at http://www.annualreviews.org/errata/pharmtox