(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 May 2003 (15.05.2003)

PCT

(10) International Publication Number
WO 03/039234 A2

(51) International Patent Classification: Not classified

(21) International Application Number: PCT/US02/35492

(22) International Filing Date:
6 November 2002 (06.11.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/338,541 6 November 2001 (06.11.2001) US
60/334,248 28 November 2001 (28.11.2001) US

(71) Applicants and
(72) Inventors: **PICKAR, David** [US/US]; 4915 Dorset Avenue, Chevy Chase, MD 20815 (US). **HIDARY, Jack, D.** [US/US]; 6500 Seven Locks Road, Suite 220, Cabin John, MD 20818 (US). **GRAY, Elizabeth** [US/US]; 6 Tomlinson Court, Cabin John, MD 20818 (US).

(74) Agents: **GATTO, James, G.** et al.; Mintz Levin Cohn Ferris Glovsky & Popeo, 12010 Sunset Hills Road, Suite 900, Reston, VA 20129 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/039234 A2

(54) Title: PHARMACOGENOMICS-BASED SYSTEM FOR CLINICAL APPLICATIONS

(57) Abstract: The invention relates to computer systems and methods of analyzing an association between patient genotypes and adverse drug phenotypes for providing personalized medical treatment, clinical trial design recommendation, and clinical management. According to one embodiment, the invention analyzes genotypes, clinical phenotypes, and clinical trial requirements to provide recommendations for conducting various phases of clinical trial process. According to another embodiment, the invention analyzes an association between patient genotypes and adverse drug phenotypes for providing personalized medical advice based on a patients personal genetic make-up.

PHARMACOGENOMICS-BASED SYSTEM FOR CLINICAL APPLICATIONS

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims priority from U. S. provisional patent application serial number 60/338,541, filed on November 6, 2001, and serial number 60/334,248, filed on November 28, 2001, each of which is incorporated by reference in its entirety.

FIELD OF THE INVENTION

The present invention relates to computer systems and methods of analyzing an association between patient genotypes and adverse drug phenotypes for providing personalized medical treatment, clinical trial design recommendation, and clinical management.

BACKGROUND OF THE INVENTION

Phramacogenomics includes identifying gene variants that influence clinical responses to drug and other treatments. This growing area of medicine enables more individualized, science-based treatment decisions. Other aspects of pharmacogenomics include predicting drug response (efficacy) and limiting side effect profiles. The ability to better predict drug response would allow individualized pharmacotherapy that could increase the chance of selecting an optimal drug for each patient and could offer savings in both time and cost of care, and substantially improve a patient's long-term prognosis.

The pharmacogenomic process includes understanding the mechanisms of action of the drug in question, identifying candidate genes based on their involvement in the mechanism of

action for the drug or illness risk factor, identifying gene variants, and determining the association of gene variants with findings from clinical trials. A drawback of existing systems for use in clinical trials is the lack of bioinformatics tools that enables efficient use of pharmacogenomics in clinical trials. Another drawback is that the existing systems lack methodologies that establish individual patient genotypes, including genome wide candidate gene and single nucleotide polymorphisms (SNP's) and detailed clinical information in a unified database to enable the clinical trial development process.

Pharmacogenomics is particularly useful in unraveling genetic bases of "complex" disorders (e.g. hypertension, diabetes, most psychiatric disorders and many cancers) as well as infectious diseases (e.g. AIDS). Complex disorders are diseases without a simple genetic inheritance, but rather those in which genetic factors effect risk phenotype (clinical manifestation), including severity and outcome, and response to pharmacotherapy. The utilization of genetic information in association with the clinical trial process would enable genetically homogenous and targeted clinical trial populations, thereby improving the "signal to noise" ratio. The value of targeted patient populations, selected by genotype of candidate genes derived from a known genomic drug mechanism pathway analysis will enhance efficiency and success rate and enable cost saving. Another drawback in the existing systems is that they lack a bioinformatics system for clinical applications that utilizes genetically selected or targeted patient populations for establishing a pharmacogenomic foundation.

The drug discovery process involves screening large number of compounds for identification of therapeutic targets. It is estimated that 2 of 5,000 compounds identified from the drug discovery process eventually reach the clinical market. Once a lead drug candidate is chosen for clinical development, the clinical trial process involves a drug regulatory agency's

(e.g., Food and Drug Administration or FDA in USA) oversight for various Phases (e.g., Phases I – III of clinical trial as per FDA's guidelines in USA). However, the existing systems lack bioinformatics features for pharmacogenomics that can examine all Phases of the Drug Development Life Cycle and provide solutions or recommendations to clinicians.

A specific application of pharmacogenomics includes identifying individuals who are at a risk of adverse events to a drug (in contrast to genetic identification of high therapeutic responders) during post-market surveillance (Phase N) period of the drug life cycle. A key element needed to provide a useful database relating to adverse events is an explicit and consistent definition of adverse event phenotype and polymorphic candidate genes based on understanding the pathways involved in the pathophysiology of the event or based on empirical observation and report without a priori hypothesis. Genetic factors related to individual differences in drug metabolism have long been recognized to affect pharmacokinetics, a key element in tolerability, optimal dose finding and other aspects of pharmacotherapeutics. Thus, genetic factors related to drug metabolism are relevant from early drug development throughout the entire drug life cycle. Therefore, another drawback in the existing systems is a lack of bioinformatics system for pharmacogenomic therapy which can utilize genetic factors related to drug metabolic issues.

In addition to metabolic issues, systemic drug adverse events are diverse and have a major impact on the market success of an otherwise successful therapeutic agent. These adverse affects fall under several categories for example: cardiac, liver, central nervous system (including behavior), hematopoetic and metabolic adverse events, etc. A systemic drug adverse event late in the pharmaceutical life cycle can be a sudden and limiting factor to a successful product.

Therefore, further drawbacks in the existing systems is a lack of bioinformatics system for pharmacogenomic therapy which can utilize systemic drug adverse events.

The pharmacogenomics may also involve the empirical association of numerous relatively low frequency gene variants into a "package" of genetic risk factors which together represent a major tool in the identification of "at risk" populations for a given adverse event. In this way, the small number of patients who might be at risk for even a relatively rare, but medically serious, adverse event might be identified prior to drug administration. This would substantially promote the success of a drug by limiting its adverse affects in its clinical application. However, the existing systems lack bioinformatics features for pharmacogenomic therapy which can analyze low frequency gene variants for adverse drug events.

Other drawbacks also exist.

SUMMARY OF THE INVENTION

The invention overcomes these and other drawbacks in existing systems. One aspect of the invention relates to a bioinformatics system that facilitates use of pharmacogenomics in clinical trials. Another aspect of the invention relates to linking biological information, including genomic and proteomic information, to the conduct and success of the clinical trial process for therapeutic agents.

In one embodiment, the present invention provides an effective system to aid in protocol design, operation, and recommendations for various phases clinical trials which incorporate pharmacogenomic principles and methods.

In another embodiment, the invention provides the system and software to enable a user to select the category of drug to be tested (e.g., antidepressant, anti-hypertensive agents), the specific mechanism of the drug in question within the drug category (e.g., serotonin reuptake

inhibitor antidepressant, ACE inhibitor antihypertensive, etc.), and receive, in an organized format, genetic information (*eg.* gene variants, SNP's, molecular markers, protein markers) including their allelic frequencies, which are related to the mechanism of action and/or have been reported to be associated with outcome measures of the drug under investigation.

In yet another embodiment, the invention further provides for ongoing patient selection balance. This involves maintaining balanced treatment "arms," involving patients with specific genotypes, wherein the system ensures sufficient statistical power needed for hypothesis testing.

In a further embodiment, the invention provides for an individual patient's clinical outcome (based on data from the clinical trial) to be merged with a personal genetic database. This combined data approach is essential for pharmacogenomic analysis of an a priori genetic hypothesis.

In an additional embodiment, the invention provides information regarding a pool of patients (identified anonymously) including detailed clinical information relating to their disease state. These patients are also genotyped for variants of candidate genes relevant to their illness or class of drug treatment for which they are candidates. In another embodiment, the invention includes whole genome-wide SNP data. In this fashion, the user of the system of the invention can effectively select patients for prospective pharmacogenetic and clinical studies.

In a further embodiment, the invention is directed to a system for controlling and utilizing genetic variants in pharmacogenetic clinical trials. The system may include a genotype database, a clinical database, an analytical computer, a clinical trial requirements database, filtering and optimization methods for clinical trial recommendation and a recommended trial database.

One aspect of the invention is directed to systems and methods of utilizing genetic variants in pharmacogenetic clinical trials by analyzing a genotype database for appropriate factors. Another aspect of the invention is directed to methods of selecting individual patients for a clinical trial by analyzing the genotypes of the patients in relation to clinical data to identify appropriate candidates. One embodiment associates a selected genotype with a clinical phenotype. Another embodiment filters genotypic and clinical phenotypic inputs based on clinical trial requirements and performs optimization of clinical trial parameters for trial recommendation.

In one embodiment, the present invention provides an effective system to aid in the identification of patients at risk for systemic drug side effects utilizing pharmacogenetic principles and methods. In another embodiment, the present invention relates to a relational database which links individualized genomics information to adverse events of therapeutic agents in medicine and provides for its organization and access.

According to one object, the invention is directed to utilizing bioinformatics technology to establish a system of databases for individual patients, including for example, their personal and genetic data, that enables the identification of genetic risk factors for adverse drug events and its application to clinical trials and market development. In one embodiment, the system provides features for establishing a database of individual patient genotypes, including genome wide and candidate gene single nucleotide polymorphisms (SNP's), and clinical information related to an adverse drug affect experienced by a patient. In another embodiment, the system creates a unified database to enable scientific understanding of risk factors for adverse events and to enable this information to be readily accessible to the clinical trial and clinical market drug development process.

According to another object, the invention provides software to enable a user to select a category of systemic drug side effects, including severity and clinical subtype, the specific mechanism of action of the drug in question within the drug category (e.g., antidepressants, antihypertensives, statins) to receive in an organized format, genetic information such as gene variants and SNP's from public databases, including their allelic frequencies, which have been associated with a given adverse event. In some embodiments, the invention allows for entry of new genetic information or individualized clinical selection criteria that is not necessarily available to the general public.

According to yet another object, the invention provides a system for screening patients in clinical trials at all stages (Phase I-N) in order to assess their risk for a specific adverse event for a specific class or individual therapeutic agent. This may enable restricting a pre-approval clinical trial to patients at lowest risk for a known side effect, thereby providing for enhanced "signal to noise ratio." It may also provide for screening of general populations for adverse event risk factors, thus strengthening the market place of a drug and minimizing the risk for adverse events in the post-market surveillance period (Phase N).

According to a further object, the invention provides information regarding a pool of patients who have experienced an adverse event to a marketed drug. Such patients may be genotyped for variants of candidate genes relevant to the side effect or class of drug treatment. This may include whole genome-wide SNP data. In this fashion, a unique individualized dataset of clinical populations who have experienced an adverse event can be matched to a corresponding dataset of genetic information.

In one embodiment, the invention is directed to a system for establishing relationships between genotype (including low frequency SNP's) and adverse events. The system may

include, for example, a genotype database, a clinical database, an analytical computer, an adverse event database, a blood bank, sequencing machines and/or clinical indications for applications of specific drugs.

In another embodiment, the invention is directed to methods of utilizing genetic variants for high throughput genotyping technologies, including, but not limited to, DNA genotyping and RNA expression "microchip arrays." The invention is further directed to methods of selecting individual patients who may be at risk for the administration of a specific drug or class or drug by analyzing the genotypes of the patients.

Other objects and features of the present invention will become apparent from the following detailed description considered in connection with the accompanying drawings that disclose embodiments of the present invention. It should be understood, however, that the drawings are designed for purposes of illustration only and not as a definition of the limits of the invention.

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1 illustrates a database system of pharmacogenomic therapy for adverse drug events according to one embodiment of the invention.

FIG. 2A illustrates a system architecture for a pharmacogenomics-based clinical trial recommendation according to one embodiment of the invention.

FIG. 2B illustrates system modules for pharmacogenomics-based clinical trial system according to one embodiment of the invention.

FIG. 2C illustrates system modules and clinical trial requirements for pharmacogenomics-based clinical trial system according to one embodiment of the invention.

FIG. 3A illustrates a pharmacogenomics-based clinical trial recommendation process according to one embodiment of the invention.

FIG.3B illustrates a process of analysis for clinical trial recommendation based on genotypic and clinical trait input according to one embodiment of the invention.

FIG. 3C illustrates a process of obtaining a clinical trial design and executing a clinical trial according to one embodiment of the invention.

FIG. 4A illustrates integration of pharmacogenomics-based clinical trial recommendation system with integrated healthcare management system according to one embodiment of the invention.

FIG. 4B illustrates the integration of a pharmacogenomics based clinical trial recommendation system, a pharmacogenomic therapy system for adverse drug events and an integrated healthcare management system according to an embodiment of the invention.

FIG. 4C illustrates a pharmacogenomic therapy process for adverse drug events according to one embodiment of the invention.

FIG. 4D illustrates a process of risk analysis for adverse drug events based on genotypic and drug phenotypic input using pharmacogenomic therapy system according to one embodiment of the invention.

FIG.5A illustrates an interface for pharmacogenomics-based clinical trial recommendation system according to one embodiment of the invention.

FIG.5B illustrates an interface for a pharmacogenomic therapy recommendation system according to one embodiment of the invention.

FIG. 5C illustrates an interface for clinical input of pharmacogenomics-based clinical trial recommendation system according to one embodiment of the invention.

FIG. 5D illustrates an interface for genetic input of pharmacogenomics-based clinical trial recommendation system according to one embodiment of the invention.

FIG. 5E illustrates an interface for inputs filtering of pharmacogenomics-based clinical trial recommendation system according to one embodiment of the invention.

FIG. 5F illustrates an interface for optimizing trial parameters of pharmacogenomics-based clinical trial recommendation system according to one embodiment of the invention.

FIG. 5G illustrates an interface for a recommendation information of pharmacogenomic therapy recommendation system according to one embodiment of the invention.

FIG. 5H illustrates an interface for a pharmacogenomic therapy system for adverse drug events according to one embodiment of the invention.

DETAILED DESCRIPTION OF THE INVENTION

The invention relates to computer systems and methods of analyzing an association between patient genotypes and adverse drug phenotypes for providing personalized medical treatment, clinical trial design recommendation, and clinical management. According to one embodiment, the invention analyzes genotypes, clinical phenotypes, and clinical trial requirements to provide recommendations for conducting various phases of clinical trial process.

According to another embodiment, the invention analyzes an association between patient genotypes and adverse drug phenotypes for providing personalized medical advice based on a patients personal genetic make-up.

As illustrated in FIG. 1, according to one embodiment of the invention, the database system of the invention may utilize genetic variants to establish a risk for adverse events using principles of pharmacogenetic science.

One or more genotype databases 52 and clinical databases 70 may be merged to establish one or more correlational databases 43, defining genotypic risks for a specific adverse event. One or more genotype databases 52 may be established through collection of biological samples (blood or other tissue) 37 analyzed using a plurality of high throughput genotyping technologies 39. A plurality of associations between contents of genotype databases 52 and clinical databases 70 may be established using an analytic computer 41 producing one or more correlational databases 43.

The one or more genotype databases 52 may include or otherwise access databases that store genotype data. Such data may include, but is not limited to, groups of individual patients who have experienced a specific adverse event to drug treatment and in whom genotype analysis for common and rare variants, including single nucleotide polymorphisms (SNP's) have been determined for specific candidate genes or have been established by a whole genome wide scan. In one embodiment, the sources for the one or more databases may include, for example, proprietary information from a user. In another embodiment, the sources for one or more databases may include, for example, public or open source information (e.g., GenBank). In yet another embodiment, the sources for one or more databases may include, for example, propreitory subscription information (e.g., Incyte Genomics Inc, Celera Genomics Corporation).

One or more clinical databases 70 shown in FIG.2 may include or otherwise access databases designed to store clinical data. Such data may include, but is not limited to, documented adverse events including the drug which incurred the adverse events, the severity and form of the adverse event (e.g., weight gain, drug-induced prolongation of QTc cardiac interval) and the outcome (cessation of drug treatment, medical care required etc.) of FDA documented adverse event profiles may be readily accessible for marketed drugs from many sources including the Physician's Desk Reference from the FDA database.

An analytic computer 41 may refer to a computer that will perform the database analyses described herein. Such a computer may be, for example, a personal computer (e.g., Pentium chip-based), Macintosh computer, Windows-based terminal, Network Computer, wireless device, workstation, mainframe computer, or other computing device. The computer may include, for example, Windows oriented platforms and include conventional software for supporting a display screen, a keyboard, a memory, a processor and input/output device (e.g., mouse). In some embodiments a plurality of analytic computers 41 may be used. In some embodiments, the plurality of computers may be connected as clusters and may be used for parallel processing.

One or more correlational databases 43 may include admixtures of clinical phenotype and genotypic data such that one or more patients may be rapidly selected on the basis of either clinical or genyotypic data to serve the needs of application risk to technologies as part of clinical application (e.g., DNA microarray).

Biological sample collection facility 37 may include a storage means in which whole blood or other tissues are received from patients who enter the database. This facility may allow

for the extraction of DNA of leukocytes, immortalization of cell lines for future DNA extraction or the maintenance of tissue for RNA expression studies.

In one embodiment, the genotyping devices 39 may include one or more analytic machines, for example, which provide for high throughput genotyping for individual candidate genes, including "deep sequencing" of large populations for low frequency single nucleotide polymorphisms or other variants. In another embodiment, the genotyping devices 39 may include a plurality of sequencing machines. In some embodiments, high throughput sequencing and genotyping may be acquired through industrial vendors (e.g., Applied Biosystems, Sequenom, Affymetrix) utilizing proprietary technology.

According to one aspect of the invention, as illustrated in FIG. 2A, a clinical trial recommendation (CTR) system 44 may include a pharmacogenomic analysis system 48 that may be used to perform genomic analysis (e.g., associating genotype with phenotype, nucleotide sequence comparison, pattern matching) and proteomic analysis (e.g., protein sequence matching, three dimensional modeling). The CTR system 44 may access and retrieve genotypic data from a genotypic database 52 and clinical data from a clinical database 70.

In one embodiment, the CTR system 44 may permit the utilization of the genotype data to carry out, design and monitor clinical trials. The genotypic database 52 may refer to databases designed to store the genotype data. Such data may include, but are not limited to, data associated with groups of individuals patients in whom genotype analysis for common and rare variants, including single nucleotide polymorphisms, have been determined for distinct candidate genes. This data may also include genome-wide SNP maps for individual patients. The genotypic database 52 may include or otherwise access expressed sequence information from an EST (Expressed Sequence Tag) database 54, microarray data from an array database 56, and/or

candidate gene data from a candidate gene database 58. The genotypic database 52 may also include or otherwise access genetic sequence (eg. nucleotide sequence, peptide sequence) from sequence bank 68. This sequence bank 68 may be able to store a large volume of genetic data including terra bytes and peta bytes of data. In one embodiment sequence bank 68 may directly access sequence data from genetic sequencing devices. In addtion, the genotypic database 52 may be coupled to other databases including map database 60, open source database 62, publications database 64, and/or user input database 66. Map database 60 may store, for example, information on genetic, physical and transcriptome maps of human and other organisms. Open source databases 62 may include, for example, public databases such as GenBank and SwissProt. The publications 64 database may include, for example, various publications including genomics, proteomics, and clinical trials. User Input database 66 may store any information specified by clinical user. The genotypic database 52 may also be coupled to proprietary databases such as, for example, Celera genomic database (not shown in figure).

The clinical database 70 may include clinical data such as, but not limited to, diagnoses confirmed by standardized assessment tools, confirmed tissue (e.g., tumor) leading to a specific disease diagnosis, illness severity, outcome for illness or syndrome, response to prior drug treatment, family and clinical genetic history, and other elements which contribute to a clinical phenotype and are associated with specific genotypes.

The clinical database 70 may include or otherwise access patient information database 76, mode of action database 72, and/or drug information database 74. Patient information database 76 may include, for example, patient information including medical history, demographical and biographical information (eg. age, sex). The mode of action database 72 may include information regarding drug mechanisms. In some embodiments, the mode of action

database 72 may include information on partial understanding of a drug mechanism for example. In other embodiments, the mode of action database 72 may provide drug mechanisms which are speculative for example. The drug information database 74 may, for example, include a list of manufacturers of a drug, dosage information, and results of previous study.

According to one embodiment, the pharmacogenomics based CTR system 44 may include a recommended trial database (not shown in Figures). The recommendation trial database may include an admixture of clinical phenotype and genotypic data such that a patient, or group of patients, may be rapidly selected on the basis of either clinical or genotypic data to serve the needs of a given clinical trial. In this fashion, a unique database may be applied to a distinct clinical trial.

According to another embodiment, the pharmacogenomics based CTR system 44 may access data (e.g., patient blood group, patient DNA source) from a blood bank (not shown in figure). The blood bank may have a storage facility in which whole blood or other tissues are received from patients who enter the database. This facility may allow, for example, the extraction of DNA of leukocytes, immortalization of cell lines for future DNA extraction or the maintenance of tissue for RNA expression studies.

According to yet another embodiment, the CTR system 44 may be coupled to a plurality of sequencing machines (not shown in figures). The sequencing machines may access biological samples of the blood bank. The sequencing machines may include analytic machines which provide for high throughput genotyping for individual candidate genes, including deep sequencing for rarely occurring single nucleotide polymorphisms or other variants.

According to another aspect of the invention, the pharmacogenomics based CTR system 44 may include a clinical trial requirements database 78. Clinical trial requirements database 78

database may include, for example, one or more inclusion and exclusion criteria for a plurality of clinical protocols. This criteria may include, for example, diagnosis, gender, age, illness severity, prior treatments, etc. In one embodiment, the clinical trial requirements database 78 may include or otherwise access FDA guidelines data.

According to yet another aspect of the invention, the pharmacogenomics based CTR system 44 may be accessed by authorized users of contract research organizations (CROs) who are involved in administering clinical trials.

According to one embodiment of the invention, pharmacogenomic therapeutic system 300 may be coupled to CTR system 44. The CTR system 44 may access therapeutic information from one or more pharmacogenomic therapeutic system 300 databases (not show in figures). In some embodiments, both pharmacogenomic therapeutic system 300 and CTR system 44 may access therapy requirements from therapy requirements database 302.

According to one embodiment of the invention, as illustrated in FIG. 2B, the CTR system 44 may include a plurality of modules for pharmacogenomics based clinical trial system. One or more genetic analysis modules 81 may be able to perform genetic analysis such as, for example, DNA sequence analysis, protein sequence analysis, genetic finger printing analysis, genetic variability analysis, haplotype analysis and phylogenetic analysis. One or more phenotypic analysis modules 83 may be able to perform conventional analysis on phenotypes such as, for example, analysis of drug response, and analysis of disease progression and intensity. One or more association modules 85 may be connected to geneotypic database 52, and clinical database 70 and may be able to determine an association between genetic information in the genotypic database 52 and clinical phenotypic information in the clinical database 70 for a plurality of patients. One or more recommendation modules 87 may be connected to genotypic database 52,

clinical phenotypic database 70, and clinical trial requirement database 78 and may be able to provide clinical trial recommendations utilizing the genetic information, the clinical phenotypic information, the clinical trial requirement information and the determined association between the clinical information and the genetic information. The CTR system 44 may be able to store output of clinical trial recommendations.

According to another embodiment of the invention, as illustrated in FIG. 2B, the CTR system 44 may further include one or more clinical workflow modules 91 for monitoring workflow during clinical trial process, one or more adverse drug event modules 93 for analyzing genetic basis of adverse reaction to a plurality of drugs, one or more clinical trial management module 95 for administration of one or more aspects of one or more clinical trial phases (Phases I – IV), and one or more pharmacoeconomics modules 97 for micro- and macro-economic aspects of clinical trials including financing and budgeting.

According to one embodiment of the invention, as illustrated in FIG. 2C, the clinical trial requirements database 78 of the CTR system 44, may include or otherwise access a drug regulatory agency (eg., Food and Drug Administration or FDA in USA) requirements database 77 and a patient database 79. The drug regulatory agency requirements database 77 may include information such as regulations (e.g., FDA regulations) and guidelines for clinical trials. The patient database 79 may include a plurality of data on patients, for example, category of patients, age information, geography, health history, and personal data. The examples of category of patients may include, for example, child, elderly, sex, ethnicity, cognitively impaired individuals, or people who are economically or educationally disadvantaged. In one embodiment, the CTR system 44 may be able to relate data within the patient database 79 using data relation modules (not shown in the figure) for determining an inter-relationship between data. For example, the

CTR system 44 may be able to determine child based on age and geography (eg. state). In general, state laws define what constitutes a "child", and such definitions dictate whether or not a person can legally consent to participate in a clinical trial.

According to another embodiment of the invention, as illustrated in FIG. 2C, the CTR system 44 may also include risk factor analysis module 98, clinical trial protocol design module 99, and database update and management module 101. Risk factor analysis module 98 may be used to predict risks or adverse effects for one or more selected individuals using information from genotypic database 52, clinical database 70, and clinical trial requirement database 78. In one embodiment, the CTR system 44 may be used to predict risks or adverse effects by relating one or more genetically selected individuals for one or more clinical traits with the data in clinical trial requirements database 78. In another embodiment, the CTR system 44 may use a plurality of statistical algorithms for predicting risks or adverse effects. Clinical trial protocol design module 99 may be used to design a protocol for clinical trial. In some embodiments, the clinical trial protocol design module 99 may access a drug regulatory agency (e.g., FDA in USA) requirements database 77 for obtaining guidelines for clinical trials (e.g., FDA guidelines). In other embodiments, the clinical trial protocol design module 99 may access with genotypic database 52, clinical database 70, and clinical trial requirement database 78. In yet other embodiments, the clinical trial protocol design module 99 may utilize the information on risks or adverse effects predicted by the CTR system 44. In one embodiment, database update and management module 101 may periodically update a plurality of databases connected to the CTR system 44 with new data. In another embodiment, the CTR system 44 may maintain the plurality of databases of the invention (e.g., genotypic database, clinical database, clinical trial database) according to a plurality of user enabled set of instructions.

FIG. 3A illustrates a clinical trial recommendation process using pharmacogenomic information. Components of the pharmacogenomics-based clinical trial recommendation process may include drug mechanism analysis, target analysis, candidate gene analysis, gene variant analysis, preliminary clinical trial analysis, association analysis, filtration analysis, clinical trial requirement analysis, and optimization of clinical trial parameters. One advantage of the present invention is that it provides assistance and guidance in managing and maximizing the efficiency of the clinical process using pharmacogenomics.

As illustrated in step 4 of FIG. 3A, drug mechanisms may be identified from the mode of action database 72. The drug mechanisms included in the mode of action database 72 may provide insight into the pharmacological processes by which a drug produces its therapeutic effects. Such drug mechanisms include, for example, alterations in function of components of dopamine systems in the central nervous system in the case of antipsychotic drugs, of cardiac adrenergic systems for some classes of antihypertensive agents, or bacterial genome expression for some antibiotics. In some embodiments, the mode of action database 72 may provide information on partial understanding of a drug mechanism. In other embodiments, the mode of action database 72 may provide drug mechanisms which are speculative.

As shown in step 8 of FIG. 3A, gene targets may be identified using the CTR system 44. In one embodiment, gene targets may be included in the genotypic database 52 to provide information regarding a drug's mechanism of action and to provide the basis for pharmacogenetics clinical trials. Such targets include, for example, the D_2 dopamine receptor as a target for antipsychotic compounds or the beta adrenergic receptor for certain antihypertensive agents.

According to one embodiment, candidate genes may be included in the candidate gene database 58 to provide the link between a target (e.g., receptor, enzyme) and its genetic control of target function and production. These candidate genes may be identified from the database in step 12.

According to another embodiment, gene variants may be included in the database to provide the genetic basis for pharmacogenetics studies. For example, the gene that codes for the D_2 receptor exists with common variants (>1% of the population) in the promoter as well as in coding regions. These variants alter an individual's production or composition of the receptor which renders this an excellent target for pharmacogenomic exploration. These gene variants may be identified in step16 from the genotypic database 52 using the CTR system 44. The gene variants may be due to, but are not limited to SNPs (Single Nucleotide Polymorphisms), variation in candidate genes, variation in number of nucleotide repeats (eg. simple sequence repeats), variation in length of nucleotide repeats, RFLPs (Restriction Fragment Length Polymorphisms), variation in protein sequences and variation in protein structures.

According to yet another embodiment, as shown in step 20, clinical trial inputs may be identified from clinical trial database 70. The clinical trial inputs may include information on one or more clinical phenotypes (e.g., mild cognitive impairment).

According to additional embodiment, an association may be established as shown in step 24 between one or more gene variants and one or more phenotypes. Once the association is determined through association analysis as shown in step 24, *a priori* hypothesis testing in further clinical trials can be accomplished. According to one embodiment of the invention, the association may be determined using a plurality of statistical methods. In one example, pearson's correlation is used to determined the association between a genotype and clinical phenotype.

According to further embodiment, the CTR system 44 may present associations between genetic information and clinical information and associated genotypes and phenotypes using a plurality of presentation tools in graphical user interface (not shown in FIG. 3A). In one embodiment, as shown in step 28, these associations may be filtered using pre-determined statistical significance or threshold values known to one skilled in the art. In another embodiment, the information may be filtered based on genes or phenotypes. For example, a user may be interested in a particular gene selected from several genes showing association for a clinical trait. In this case, the user may be able to select one or more preferred genes and filter out the genes and other information related to the genes which are not preferred.

According to one embodiment, the CTR system 44 may be used to obtain a plurality of clinical trial requirements as shown in step 32 . The clinical trial requirements may include, for example, Food and Drug Administration guidelines for various phases of clinical trials. The clinical trial requirements may correspond to, for example, diagnosis, gender, age, illness severity, and/or prior treatments of clinical patients. The CTR system 44 may be used to perform optimization of the plurality of clinical trial requirements using the genotypic and the phenotypic input as shown in step 36. For example, the CTR system 44 may be used to optimize the clinical trial requirements for children in the age group of 10-14 since the clinical trial requirements may be dependent on risk factors in a developmental stage or age of the clinical patients. The CTR system 44 may provide clinical trial recommendation, as shown in step 40, utilizing the results of the optimization.

According to one embodiment of the invention, a process for determining a clinical trial recommendation based on genotypic and clinical trait input is illustrated in FIG. 3B. For example, in a clinical study, a plurality of genotypes 114, a plurality of clinical traits 116, and a

plurality of clinical trial requirements 78 may be analyzed at step 100 using one or more analytical processors. The clinical trait may include any clinical phenotype such as response to drug, dosage of drug, patient age etc. In this analysis, individuals having similar genotypes and similar clinical traits may be selected and grouped together. For example, one or more selective genotypes may be associated with one or more selective phenotypes. The selected genotypes or clinical traits may be included or excluded depending on the nature of the clinical study. In one embodiment, genotypes with high similarity may be included for a clinical study. In another embodiment, dissimilar genotypes may be included for a clinical study. In yet another embodiment, genotypes may be randomly chosen to have genetic balance, and included in a clinical study. In a further embodiment, the invention provides for ongoing patient selection balance. This may involve maintaining balanced treatment "arms," involving patients with specific genotypes, thereby ensuring sufficient statistical power needed for hypothesis testing.

The selected genotypes and clinical traits may be analyzed at step 100 with the plurality of clinical trial requirements 78 of a given clinical study. If the selected genotypes and clinical traits meet the clinical trial requirements, they may be validated at step 108 against the plurality of clinical trial requirements of individual phases (e.g., Phase III) of a clinical trial. The trials may be recommended based on output of analysis. If selected genotypes and clinical traits do not meet the clinical trial requirements, the results may be stored at step 112 and may be used for further analysis.

According to another embodiment of the invention, the process of obtaining a clinical trial design and executing a clinical trial are illustrated in FIG. 3C. For example, in a clinical study, a plurality of genotypes 114, a plurality of clinical traits 116, and a plurality of clinical trial requirements 78 may be analyzed at step 100 using one or more analytical processors. In

this analysis, individuals having similar genotypes and similar clinical traits may be selected and grouped together as shown in step 113. Clinical trial protocol may be designed for a trial involving selected individuals as shown in step 118. The protocol may consider a plurality of parameters including, for example, risk or adverse drug effect information for selected individual, patient category, sex, age, geography, health history and personal data. The protocol may be submitted electronically to a group of authorized individuals (eg. Institutional Review Board) for review and approval (not shown in FIG. 3C). In some embodiments, a user may authorize a group of individuals to access one or more of the features of the system 44 or one or more of the features connected to the system 44 as part of review and approval of the protocol. After obtaining the approval for the protocol as shown in step 118, the protocol may be executed using the CTR system 44 as shown in step 119.

According to another embodiment, as illustrated in Fig. 4A, the CTR system 44 may be integrated with an integrated health care management system 120. The integrated healthcare management system 120 may, for example, interact with one or more organizations for managed care systems (eg. PPO, HMO), and a plurality of healthcare users 124 such as healthcare managers, paramedical specialists and physicians. In some embodiments, the healthcare users 124 may have access to a CTR system 44.

As illustrated in FIG. 4B, according to one embodiment of the invention, a pharmacogenomics based CTR system 44, a pharmacogenomic therapeutic system 300 and an integrated healthcare management system 120 may include means to interface and communicate with each other. These systems may have means to access and retrieve genotypic data from one or more genotypic databases 52 and clinical data from one or more clinical databases 70. As illustrated in FIG. 4B, the pharmacogenomic therapeutic system 300 may access genetic data

from one or more genotypic databases 52, clinical data including adverse drug event data from one or more adverse event databases 304 through one or more clinical databases 70 and patient data from one or more patient databases 76. In some embodiments, the pharmacogenomic therapeutic system 300 may access adverse drug event data directly from one or more adverse event databases 304. Phenotypic characterization of the adverse event may be included in the database to provide insight into the pharmacogenomic processes by which a drug may produce a specific adverse event. Such adverse events may be characterized initially by the affected physiological system (e.g., cardiac, behavioral, endocrine, etc.).

According to one embodiment, the one or more clinical databases 70 may access one or more adverse event databases 304 and one or more drug information databases 74. According to another embodiment, the pharmacogenomic therapeutic system 300 may enable patients 316, a plurality of healthcare users 308 such as healthcare managers, paramedical specialists and/or physicians to access a patient database 76. In some embodiments, this access may be restricted by a plurality of authorization means. According to yet another embodiment, the plurality of healthcare users 308 may access pharmacogenomic therapeutic system and analyze genetic data, adverse event data and patient data for providing personalized medicines.

According to another aspect of the invention, the pharmacogenomic therapy system 300 may be integrated with an integrated health care management system 120. The integrated healthcare management system may refer to a system that interacts with one or more organizations for managed care systems (e.g., PPO, HMO), and the plurality of healthcare users 308. The healthcare users 308 may also access CTR system 44.

In one embodiment, the present invention may permit the utilization of genetic data to gain molecular understanding of adverse events. In another embodiment, the present invention may enable the user to access clinical information about the individual patient's adverse events from a clinical database 70 in relation to that person's individual genomic information. The resultant analyzed database may provide the user with individual patient and/or group information related to an adverse event to a specific drug category (or drug) regarding the genetic associations with the adverse event in relation to genotypes. The system 300 of the invention may provide specific information regarding genotypic relationships between adverse events and specific drug treatments. As such it will be utilized by, for example, pharmaceutical, contract research organizations, site management organizations during clinical development of a new therapeutic agent.

In another embodiment, the invention may allow discovery programs from biopharmaceutical companies to explore genetic relationships to adverse events by providing biological and clinical material from patients in the database who have experienced the adverse event in question. In this way "deep sequencing" efforts (sequencing of large numbers of subjects, e.g., 2500) may be accomplished in order to identify rare SNP's or other variants related to the adverse event. This information may be utilized by the CTR system 44 to establish a database that could identify new genetic based "targets" for drug discovery programs. The risk of adverse events may in this way be minimized early in the small molecule clinical development process.

In yet another embodiment, the present invention may enable the estimation of risk for an adverse event in patients who might be suitable for the therapeutic administration of an approved drug. The user first may enter the drug category (e.g., antidepressant, antihypertensive,

antibiotic) and specific therapeutic agent (e.g., fluoxetine, atenelol, Cipro, etc.) for which the patient is a candidate as part of his or her therapeutic regimen. The pharmacogenomic therapeutic system 300 may provide information regarding adverse events and their known association with genetic risk factors for specific drugs or drug categories. The user may also enter the category of adverse event (e.g., cardiac, behavioral) and receive genetic risk factors for the adverse event that may extend across therapeutic agents. The user may then apply a DNA array or other genotyping technologies to biological material from an individual patient in order to gain an estimate of the risk for the adverse event.

FIG. 4C illustrates, according to one embodiment of the invention, a pharmacogenomic therapy process for adverse drug events using pharmacogenomic information. Components of the pharmacogenomics based therapy may include for example: drug information analysis, adverse drug event analysis, drug mechanism analysis, gene target analysis, candidate gene analysis, gene variant analysis, preliminary clinical trial analysis, association analysis, validation analysis for association, and/or prescription recommendation analysis. Information on one or more of drugs may be obtained as shown in step 2 from one or more drug information databases 74. Similarly one or more adverse drug events may be obtained from one or more adverse event databases 304, as shown in step 3, and adverse events of one or more drugs may be identified and analyzed. Adverse events might include, for example, hypotensive reactions or heart rhythm irregularities (e.g., QTc prolongation), drug-induced diabetes (endocrine), psychotic reactions (behavioral), etc. In some instances adverse events may involve multiple physiological systems with multiple clinical manifestations.

As shown in step 4 of FIG. 4C, drug mechanisms may be identified from one or more mode of action databases 72. The drug mechanisms included in the one or more mode of action

databases 72 may provide insight into the pharmacological processes by which a drug produces its therapeutic events. Such drug mechanisms include, for example, information on alterations in function of components of dopamine systems in the central nervous system in the case of antipsychotic drugs, cardiac adrenergic systems for some classes of antihypertensive agents or bacterial genome expression for some antibiotics. In some embodiments, partial understanding of a drug mechanism may be obtained. In other embodiments, information on drug mechanisms which are speculative may be obtained. In yet other embodiments, drug category and information regarding the therapeutic mechanism of action (and known adverse events) of the drug in question may be obtained for the purpose of identifying genetic targets related to the causation of the adverse event. Examples of drug categorization may include, for example, thioridazine, an antipsychotic (a.k.a. neuroleptic) agent within the Phenothiazine chemical group; or the antihypertensive agent, atenolol, a representative of the Benzeneacetamide chemical group, belonging to the therapeutic class of B1-adrenergic blockers; or Cipro (ciprofloxacin), a broad spectrum antibiotic of the fluroquinolone chemical group.

As shown in step 8 of FIG. 4C, the present invention enables one to identify gene targets. In one embodiment, gene targets may be included in one or more genotypic databases 52 to provide information regarding a drug's mechanism of action and to provide a basis for pharmacogenomic therapy. In another embodiment, the targets may be included in the one or more genotypic databases 52 to provide information regarding both the drug's mechanism of action and pathophysiological pathway for an adverse event. This might provide the basis for application of pharmacogenetics for risk identification. Such targets may include, for example, striatal D2 receptors for extrapyramidal side effects of antipsychotic drugs or cytochrome P450

for pharmacokinetic variability of the numerous drugs which are metabolized through the cytochrome P450 system.

According to one embodiment, candidate genes of the invention may provide a link between the target (e.g., receptor, enzyme) and genetic control of the target's function and production. These candidate genes may be identified in step 12 from one or more candidate gene databases 58 of the present invention.

According to another embodiment, the invention may include or otherwise access information on gene variants and information on the genetic basis for pharmacogenetics studies. For example, the gene that codes for the D_2 receptor exists with common variants (>1% of the population) in the promoter as well as coding regions. These variants alter an individual's production or composition of the receptor which renders this an excellent target for pharmacogenomic exploration. Common gene variants of specific enzymes of the P450 cytochrome system may enable characterization of patients into three distinct metabolizing patterns: rapid, intermediate and slow. These gene variants may be identified in step16 from the genotypic database 52 using the pharmacogenomic therapeutic system 300. The gene variants may be due to, but not limited to, SNPs (Single Nucleotide Polymorphisms), variations in candidate genes, variations in number of nucleotide repeats (e.g., simple sequence repeats), variations in length of nucleotide repeats, RFLPs (Restriction Fragment Length Polymorphisms), variations in protein sequences and/or variations in protein structures. In some embodiments gene variants may be scanned all over the entire genome. A genome wide scan may enable the search for genetic susceptibility to disease or adverse event without initial focus placed on a specific candidate gene. Scientists using rapidly emerging "haplotype" maps of the genome may more readily be able to "scan" throughout the entire genome for linkages and associations

between phenotype and genotype. Haplotypes are ancestral segments of chromosomes that contain many SNP's inherited together as a set or a block enabling easier, faster and less expensive ways to find disease or adverse event causing or predisposing genes which may be characteristic of individual patients. Genome wide scans may be performed on data in the genotypic database 52 for enabling the assembly of a detailed haplotype (SNP block) profile for the adverse event.

According one embodiment, the CTR system 44 may obtain clinical trial information as shown in step 20 and perform association analysis using the genotypic and the phenotypic input. According to one aspect of the invention, an association may be established in step 24 between one or more gene variants and one or more phenotypes (e.g., adverse response to drug, drug mechanisms). Once the association is determined through association analysis in step 24 of the system components, *a priori* hypothesis testing in further clinical trials can be accomplished. According to one embodiment of the invention, the association may be determined using a plurality of statistical methods. In one example, a pearson's correlation may be used to determine the association between a genotype and clinical phenotype. According to another embodiment of the invention, the associated patient genotype and drug phenotype may be validated in step 29 using one or more statistical methods known to one skilled in the art.

According to another embodiment, clinical and genetic data may be admixed into the one or more correlation databases 43. Once a relationship is established between one or more genotypes and one or more adverse events through association analysis of the database components, the information may be used to develop screening or other clinical monitoring techniques to identify patients who might be at risk for experiencing the adverse event. Numerous SNP's and other candidate gene variants may be assembled onto a DNA microarray

"chip" or other technologies which may enable rapid multiple genotyping for one or more individual patients, thereby creating a clinical efficient and validated method for establishing pharmacogenetics risk for an adverse event. This methodology may then be applied broadly as a clinical screening tool for patient populations.

According to one aspect of the invention, the CTR system 44 may be able to bring genetic information and clinical information of associated genotypes and phenotypes. These associations may be filtered using a pre-determined statistical significance or threshold value. In one embodiment, the information may be filtered based on genes. For example, a user may be interested in a particular gene selected from several genes showing association for a clinical trait. In this case, the user may be able to select one or more preferred genes and filter out the genes and the information related to the genes which are not preferred. In another embodiment, the information may be filtered based on one or more preferred phenotypes. In yet another embodiment, the information may be filtered based on one or more preferred associations between one or more genotypes and one or more phenotypes. According to another aspect of the invention, information on associations and validated associations may be used for further analysis for recommending prescriptions in step 33.

According to one aspect of the invention, the process of analysis for a therapy based on genotypic and drug related phenotypic information and recommending a drug as prescription is illustrated in FIG. 4D. In one embodiment, a plurality of genotypes 114, and a plurality of drug related phenotypes 115, may be analyzed using one or more analytical processors. The drug related phenotype may refer to traits such as response to drug, dosage of drug, adverse event of drug, severity of adverse events, etc. In this analysis, individuals having similar genotypes and similar drug related phenotypes may be selected and grouped together. One or more selective

genotypes may be associated with one or more selective phenotypes. Means for inclusion and exclusion of selected genotypes and phenotypes may be provided. These inclusions and exclusions may depend on nature of a therapeutic analysis. In one embodiment, genotypes with high similarity may be included for a therapeutic analysis. In another embodiment, genotypes may be randomly chosen to have genetic balance, and included in a therapeutic study. In a further embodiment, the invention provides for ongoing patient selection balance. This involves maintaining balanced treatment "arms," involving patients with specific genotypes, wherein the system ensures sufficient statistical power needed for hypothesis testing.

In one embodiment, the selected genotypes and drug related phenotypes may be analyzed with the patient related information (e.g., age of patient, health history of patient, etc.). In another embodiment, the selected genotypes and drug related phenotypes may be analyzed with the therapy requirements. Therapy requirements may include, for example, classes of medication, choice of specific medication, etc.. In yet another embodiment, the selected genotypes and drug related phenotypes are analyzed with clinical trial data including plurality of clinical trial requirements of individual phase (e.g., Phase III) of a clinical trial.

According to one aspect of the invention, the pharmacogenomnic therapy system 300 may obtain therapy requirements data from the plurality of therapy requirements database 302, as shown in step 116. In some embodiments, the pharmacogenomnic therapy system 300 may include means to select genotypes and analyze drug phenotypes with the plurality of therapy requirements. Associations among the selected genotypes, the drug phenotypes and the therapy requirements may be determined using one or more algorithmic methods (e.g., hidden-markov based analysis, artificial intelligence and neural network, etc.,).The associated genotypes, phenotypes and therapy requirements may be further analyzed for risk for adverse drug events

using one or more pre-determined formulas or algorithms, as shown in step 301. If there is no risk for adverse events and the selected drugs are suitable for prescription, the analysis may be validated using a plurality of statistical validation models known to one skilled in the art. In some embodiments, the analysis results may be validated against the plurality of clinical trial requirements of an individual phase (e.g., Phase III) of a clinical trial. In other embodiments, the invention provides a system for screening patients in clinical trials at all stages (Phase I-N) in order to assess their risk for a specific adverse event for a specific class or individual therapeutic agent. This may enable restricting a pre-approval clinical trial to patients at the lowest risk for a known side effect, thereby providing for enhanced "signal to noise ratio." It may also provide for screening of general populations for adverse event risk factors, thus strengthening the market place of a drug and minimizing the risk for adverse events in the post-market surveillance period (Phase N).

According to another aspect of the invention, as shown in step 303 of Figure 4D, the system 300 may determine whether or not a selected drug is suitable for prescription for one or more diseases or disorders of a selected individual based on the results of the analysis of adverse drug events (discussed above). If the selected drug is not suitable for prescription, the results of the analysis may be stored, as shown in FIG. 4D. If the selected drug is suitable for prescription, the system may perform additional validation or secondary validation of this prescription using one or more user selectable validation models which are not used in previous analysis, as shown in an optional step 305. In some embodiments, the system 300 may enable a user to recommend prescription for the selected individual based on one or more of the analysis procedures (discussed above) for adverse drug events, as shown in step 307.

FIG. 5A illustrates a user interface 130 for CTR system of FIG. 4, item 44, according to one embodiment of the invention. The user interface 130 may have a plurality of icons (e.g., clickable buttons) for managing clinical data 134, managing genomic data 138, defining clinical trial 142, recommending clinical trial 144 and managing clinical trial 148. Manage clinical data button 134 may be used to access database management features of pharmaceutical, patient, and other clinical phenotypic databases, for example, in the CTR system 44. Clinical database management features may support entry and editing of data in the clinical databases. The relationships among data and databases may also be managed using these features. In one embodiment, the clinical database management features may include user intervened data update features. In another embodiment, the clinical database may be managed and updated automatically without user intervention. In some embodiments, the clinical database management features may include a plurality of frames preferably in a graphical user interface for performing database maintenance functions.

Manage genome data button 138 may be used to access genetic data (e.g., nucleotide sequence, protein sequence, protein structural data, protein functional data, genome map, etc.) and publications and reports relevant to genetic data of both proprietary and public databases, for example. Furthermore, the user may operate genome database management features through manage genome data button 138 for entering and editing of data in the genomic or genetic databases of the system 44. For example, the user may manage the relationships among genetic data and databases. In one embodiment, the genome database management features may include user intervened data update features. In another embodiment, the genome database may be managed and updated automatically without user intervention. In some embodiments, the

genome database management features may include plurality of frames preferably in graphical user interface for performing database maintenance functions.

A clinical trial may be defined using the define clinical trial button 142. This button 142 may be used to access a plurality of frames, wherein trial information may be recorded and stored. In some embodiments, the system may have a pre-determined format for entering clinical trial information. In other embodiments, the user may be able to create the formats. These formats may correspond to FDA requirements for clinical trials. In one embodiment, the present invention provides an effective system to aid in protocol design, operation, and recommendation for Phase I - III clinical trials and post-market surveillance that utilize pharmacogenomic principles and methods.

Manage clinical trial data button 148 may be coupled to database management features to manage data during the clinical trial. For example, trial status, diagnoses, treatments, and outcomes may be managed. According to one embodiment, clinical trial management features may support data imported from other data systems containing patient data or direct input. A plurality of import/edit screens may be used to show how the clinical trial is being managed.

According one embodiment, as illustrated in FIG. 5B, a user interface 430 may include a plurality of inputs (i.e. clickable buttons) for managing clinical data 134, managing genomic data 138, analyzing therapy requirements 442, recommending pharmacogenomic therapy 444 and managing pharmacogenomic therapy 448. Pharmacogenomic therapy requirements may be analyzed using button 442. This button 442 may enable a user to access a plurality of frames (not shown in figure), wherein information on therapy requirements of a plurality of diseases / disorders may be recorded. In some embodiments, the system may include a pre-determined format for entering therapy requirement information. In other embodiments, the user may create

the formats. These formats may correspond to requirements specified by healthcare organizations.

Manage pharmacogenomic therapy button 448 may be coupled to database management features (not shown in FIG. 5B) to manage data during the therapy. For example, the health status of patient, diagnoses, treatments, and outcomes may be managed. According to one embodiment, pharmacogenomic therapy database management features may support data import from other data systems containing patient data. A plurality of import/edit screens may be used to for pharmacogenomic therapy database management.

Clinical trial recommendation button144 or pharmacogenomic therapy recommendation button 444, may be used to view an interface 152 as illustrated in Figures 5C, 5D, 5E, and 5F and interface 452 as illustrated in FIG. 5G. The clinical trial recommendation interface 152 may have means for inputting, for example, clinical and genetic information, filtering the information and optimizing trial parameters for trial recommendation. For example, the interface 152 of Figures 5C, 5D, 5E, and 5F may include user selectable frames such as clinical input 154, genetic input 158, input filters 162 and optimize trial parameters 166 in graphical user interface. According to one embodiment, a plurality of clinical phenotypic records may be obtained, analyzed and managed using clinical input frame 154 as illustrated in FIG. 5C. The clinical input interface 154 may include a plurality of options for the user to select one or more clinical phenotypic traits or enter a clinical phenotypic trait to be used in the clinical trial. The examples of the clinical phenotypic traits may include, for example, diseases (e.g., Alzheimer), disorders (e.g., cognitive impairment), drugs (e.g., dopamine), categories of drugs (e.g., antidepressant, anti-hypertensive agents), mechanisms of drugs (e.g., serotonin reuptake inhibitor antidepressant; ACE inhibitor anti-hypertensive), etc. As illustrated in FIG. 5C, according one embodiment, the

user may enter patient ID in box 170 and retrieve individual patient data including patient phenotypic data from patient database 76. In another embodiment, the user may select a clinical phenotypic trait and analyze clinical phenotypic information of a group of patients using clinical input frame 154. For example, the user may enter disease phenotype in box 174 and retrieve disease data from the clinical database 70. The disease data may include, but are not limited to, symptoms of disease, diagnostic information and treatment information. Similarly, the user may enter drug response phenotype in box 182 and retrieve drug data from the drug information database 74.

According to another aspect of the invention, the user may input drug related information such as, for example, category of drug, mechanism of drug, etc. In one embodiment, the user may select a drug category from scroll down menu 186. In another embodiment, the user may select a drug mechanism using scroll down menu 190. The system 44 may obtain the information related to selected drug category or drug mechanism from the drug database 74. In addition, the user may stratify the selected clinical phenotypic traits based on a plurality of statistical models known in the art for stratification. The user may use scroll down menu 178 for selecting a statistical model for stratification. In one embodiment, the statistical model for stratification may correspond to phenotypic correlation of individuals. In another embodiment, the statistical model for stratification may correspond to chi-square methodology for grouping individuals. Stratification of individuals based on their clinical phenotypic traits may enable clinicians to target clinical study to a group of individuals with similar clinical phenotype.

According to one embodiment, the user may enter information regarding genetic markers that pertain to biological mechanism of a specific drug undergoing clinical trial and the CTR system 44 may balance distributions of genotypes among study populations undergoing specific

clinical trials. Thus, the invention provides the ability to monitor the composition of clinical trial populations during the conduct for the clinical trial.

According to one embodiment, the user may select individual patients who are suitable for a clinical trial on the basis of already performed genotypes. For example, the user may first enter the category of drug in a trial (e.g., antidepressant, anti-epileptic, etc), may next select a specific pathway of its mechanisms (e.g., serotonin reuptake blockage) or describe a pathway not yet existing in the data base, and finally may identify known candidate genes and their variants in the database which could pertain to the drugs therapeutic action on the basis of information.

The genetic input of clinical trial recommendation is illustrated in Figure 5D. According to one aspect of the present invention, the user may select one or more genetic input from the genetic input frame 158. In one embodiment, the user may enter a gene identification number or a gene name in box 194 and obtain a plurality of information related to the specified gene from the genotypic database 52. In another embodiment, the user may enter more than one gene or multiple genes in box 198 and obtain information related to multiple genes from genotypic database 52. The information on multiple genes may correspond to clinical studies of complex diseases since the complex diseases are known to be controlled by multiple genes. In yet another embodiment, the user may select a plurality of database sources for obtaining genetic data. The genetic data may include, but are not limited to, SNP (single nucleotide polymorphism), EST (Expressed Sequence Tags), protein data, and candidate genes. These data may be obtained from one or more databases such as Seq. Bank 68, EST DB 54, and candidate gene DB 58 of system 44. The genetic input frame 158 may have a link to a genetic analysis system 216, wherein the genetic analysis system 216 enables the user to perform genomic (eg. sequence matching and gene identification, gene expression analysis, genotype analysis) and proteomic (protein

identification, predicting protein structure, predicting protein-protein interactions) analysis. The genetic input frame 158 may also have link to a statistical analysis system 220, wherein, the statistical analysis system 220 enables the user to analyze genetic data using a plurality of statistical or mathematical methods (eg. principal component method for gene expression, regression methods for genotype association, Hidden-Markov methods for sequence matching). The statistical analysis system may enable the user to group or stratify individuals based on a plurality of genetic similarities. In some embodiments, the selected genes may be allelic variants. The allele frequency selected genes may be displayed in box 202.

According to another aspect of the invention, as illustrated in FIG. 5E, the user may associate the selected genetic inputs with the selected clinical phenotypic inputs. These associations may be determined using one or more of statistical tests. For example, the user may perform correlation test as shown in box 224 of FIG. 5E. The association may be performed between one or more genes including allelic variants and one or more clinical phenotypic traits. The user may filter the associations using a plurality threshold levels for selecting the associated samples. For example, in one embodiment, the threshold level for correlation may be selected from box 228. In some embodiments, the threshold levels may be pre-determined. Clinicians and researchers involved in clinical trial may be interested in focusing on a few genes or selecting a few genes. Similarly, they may be interested in a few aspects of information relevant to phenotypic traits. According to one embodiment of the invention as illustrated in FIG. 5E, the user may filter the selected clinical and genetic inputs and the retrieved information related to the selected clinical and genetic inputs. The genetic input may be further selected from box 232. The further selected genetic input may be displayed in box 236. Similarly, the clinical phenotypic input may be further selected from box 240 and the further selected phenotypic input may be

displayed in box 244. According to one embodiment of the invention, the user may filter the inputs using one or more filtering models. The filtering models may include parameters such as, for example, a threshold level for association between genetic input and clinical input, a threshold level for determining a similarity between the selected genetic or phenotypic input and the retrieved information from one or more databases in the system 44. According to one aspect of the invention, when the user knows which candidates are pertinent to the drug trial, the CTR system 44 may enable the choice of specific patients, that are already categorized by patterns of candidate gene variants and/or single nucleotide polymorphism (SNP) patterns. In another aspect of the invention, the CTR system 44 may enable the organizers and managers of clinical trials to establish and select pre-hoc trial populations which enable hypotheses of genetic variants as predictors of therapeutic response to be tested in an efficient and scientifically rigorous fashion.

According to another aspect of the invention, the system provides optimization features for clinical trials. As illustrated in FIG. 5F, the optimization trial parameter frame 166 may include a plurality of optimization parameters, wherein the optimization parameters correspond to a plurality of clinical trial requirements. The user may select one or more optimization parameters and perform optimization using selected clinical phenotypic inputs and genetic inputs. Since the clinical trial requirements for various phases may be different, the user may select the phase of the clinical trial from box 266. One or more protocols for clinical trials may be provided in box 270. The user may select, for example, a plurality of pre-determined inclusion/exclusion criteria from box 274. The user may also specify a clinical trial design 278. For example, the user may select clinical trial designs such as single-blind trial, double-blind trial, crossover trial and open label trial. In a single-blinded trial, the participants do not know whether they're receiving a treatment or placebo (control) until the trial is over. In a double-

blinded trial, neither the participants *nor* the researchers know who is receiving a treatment or a placebo until the trial is over. Sometimes, midway through the trial, the group receiving the treatment switches to the placebo, and vice versa, with neither group knowing which substance is which. This crossover is done to address ethical concerns about depriving one group of a possibly beneficial treatment for the duration of the trial. Crossover trial designs encourage trial participation by promising all participants access to the experimental treatment for half the trial's duration. In an open-label trial, everyone involved "sees the label" on the drug container and knows what he/she's taking.

The user may also have randomization options in box 286. The user may randomize the individuals to be involved in clinical trials. In one embodiment, the randomization may be performed within the selected individuals of similar genetic make-up. In another embodiment, the randomization may be performed within the selected individuals of similar clinical phenotypes.

According one embodiment of the invention, the system 44 may provide clinical trial recommendation based on optimization of clinical trial parameters utilizing clinical trial phenotypic input, genetic input and clinical trial requirements. In one embodiment, as illustrated in FIG. 5F, the user may run optimization and obtain recommendation of clinical trial by clicking box 290. According to another embodiment, the present invention may provide means for operating at least one phase of the clinical trial based on clinical trial recommendations.

While there are tools available to organizing information about commercial clinical trials -- cost, billing, inclusion criteria, patients screened and entered into trials --- the present invention addresses the specific need for genetic information and provides for constructing, maintaining and monitoring clinical trials on this basis. This will have operational relevance to

pharmaceutical, contract research organizations, site management organizations and clinical research specialists.

According to another aspect of the invention, the system provides optimization features for clinical trials. As illustrated in FIG. 5G, a user may select one or more therapy requirements using box 565. These therapy requirements may obtain data from therapy requirements database 302. The associated genotype and drug phenotypes may be analyzed using therapy requirements to enable pharmacogenomic therapy for plurality of patients.

In one embodiment, the associations for pharmacogenomic therapy may be validated using a plurality of validation models in box 564. These validation models may be statistical or mathematical models including but not limited to artificial intelligence and neural network methods, maximum likelihood methods, principal component methods, and combinatorial matrix algorithms. These methods may include genotypic variables, phenotypic variables and variables related to one or more therapy requirements specified by the user. In some embodiments, these methods may correspond to features or means that could convert qualitative information into quantitative variables and may include such variables in validation.

According to another embodiment, the user may obtain a plurality outputs for therapy recommendations. In one embodiment, the user may be able to select one or more presentation formats from box 568. After selecting the required elements for validation of recommendation, validation models may be run by clicking box 570.

According to another embodiment of the invention, as illustrated in FIG. 5H, an interface for the pharmacogenomic therapy recommendation for adverse drug events may include features for enabling a user to enter, for example, patient information, features for extracting and analyzing patient genetic data and patient clinical data, features for enabling a user to enter drug

information and recommend a prescription utilizing a plurality of prescription analysis models. The user may enter a patient ID in box 334 and obtain patient related information. The user may obtain specific information about a patient. Patient genetic data may be obtained by clicking one of the options in scroll down menu box 338. These options may include, but are not limited to, SNP (single Nucleotide Polymorphism) variants, candidate gene variants, simple sequence repeat variants and protein structure variants. Patient clinical data may be obtained using box 346. The examples of patient clinical data may include, for example, patient health history, age, and demographical information. In one embodiment, the user may enter one or more drugs in box 342 and retrieve adverse events of the entered drugs. The user may perform risk analysis of the entered drugs for adverse drug events. The user may store output of the analysis using item 350. Prescription recommendation analysis for pharmacogenomic therapy may be performed by selecting one or more prescription analysis models provided in 354. These models may include statistical or mathematical methods which utilize information from patient genetic data, patient phenotype data and selected drug data for predicting risk for adverse drug events. In one example, artificial intelligence and neural network model is used for prescription recommendation. In another example, principal component analysis is used for prescription recommendation. In yet another example, combinatorial matrix approach is used for prescription recommendation. In one embodiment, the user may have options for selecting one of the pre-determined statistical or mathematical models.

The present invention may enable healthcare users including physicians, researchers and clinicians to gain information regarding genetic risk factors for specific adverse events related to individual drugs across therapeutic categories and provides a mechanism of applying this information to the patient. It may also provide for the use by companies wishing to establish

proprietary diagnostic tools based on findings from internal discovery programs or specialized information related to their drugs.

While a particular embodiment of the present invention has been described, it is to be understood that modifications will be apparent to those skilled in the art without departing from the spirit of the invention. The scope of the invention, therefore, is to be determined solely by the following claims.

The invention will be better understood by reference to the following non-limiting examples.

Example #1

A pharmaceutical company may wish to bring a lead compound targeted as an antidepressant into clinical trials. The system of the invention can be used to assist in such efforts.

In this example, the compound has already passed through Phase I trials and showed no limiting adverse events in normal controls. The pharmaceutical company desires to enter this compound into Phase II trials both to establish preliminary efficacy and dose finding. The pharmaceutical company may also wish to gain information regarding how the compound might successfully establish itself within a crowded but highly lucrative therapeutic area. Towards this end, the pharmaceutical company chooses to initiate a pharmacogenomic component to the Phase II trial in order toachieve preliminary indication that a specific genotype may predict favorable response to the drug. This genotype could be used to identify prospective patients in clinical settings who would benefit from the drug administration. In this way the pharmaceutical company establishes its initial market niche.

The pharmaceutical company may utilize the system of the present invention to design, operate and monitor the pharmacogenomic clinical trial. The company may utilize the system of the invention in the following fashion:

1. The compound may be known, for example, on the basis of its activity against *in vitro* targets to belong to a class of antidepressants which inhibit the activity of the serotonin transporter, the principle neuronal mechanism for terminating the physiological effect of the neurotransmitter serotonin once it is released into the synapse (space between two adjoining neurons). This group of agents is often referred to as selective serotonin reuptake inhibitors (SSRI's). In this way, this compound enhances and regulates the brain's serotonin system.
2. It is well known that drugs of this category, while effective, do not have similarly favorable effects on every patient. Indeed, some patients experience remarkably favorable effects, while other patients remain treatment resistant. Moreover, not all SSRI's have the same therapeutic profile in individual patients. Drug choice is largely related to empirical (trial and error) experience. Establishing a genetic marker which predicts favorable response to the antidepressant could establish a market segment for this compound.
3. The pharmaceutical company may utilize the CRT system 44 to access genotypic and phenotypic information on SSRI and examine what genetic variants are known to directly relate to the pathway (mechanism of action) of this compound.
4. The CRT system 44 informs the company that there are, for example, two common variants (long and short) in the promoter region of the serotonin

transporter gene (5HTTLPR). The long variant has been reported to be associated with favorable response to an SSRI drug.

5. The system may also reveal, for example, several other gene variants, including but not limited to a functional variant in the tryptophan hydroxylase gene (TPH), and a functional variant in the promoter region of the dopamine transporter gene (DAT) which have relevance to brain function.

6. The pharmaceutical company may decide that the key gene target is the 5HTFLPR gene variant. The company has broad interest in the other gene variants as well.

7. The CRT system 44 may utilize its database to design a clinical study which balances frequencies of 5HTTLPR gene variants in study arms (placebo/active drug x 2 doses). This may include the inclusion of patients who meet clinical criteria for the antidepressant trial and also include sufficient representation of patients with each of the two genotypes.

8. The CRT system 44 may allow for genotyping for the 5HTTLPR gene and selects from the resultant patient pools (which may include referrals received from contract research organizations and site management organizations) candidates for the study which are pre-genotyped for the target gene variant.

9. Patients chosen for the study on the basis of clinical and genotypic characteristics for the 5HTTLPR gene may also be genotyped for a selected group of exploratory gene variants (e.g., DAT).

10. The CRT system 44 may provide sufficient statistical power (distribution of 5HTTLPR long and short genotypes) regarding the potential use of this genotype

as a predictor of drug response. The CRT system 44 may also enable a user in the pharmaceutical company to perform exploratory examination of additional gene variants.

11. The findings from this Phase II study may indicate a statistically or near statistically significant "signal" supporting the long variant of the 5HTTLPR as predictive of favorable drug response.

12. The pharmaceutical company may choose to advance clinical investigation for this compound into Phase III, using the now established dose, and may generate an *a priori* hypothesis regarding favorable drug response and the long form of the 5HTTLPR gene. This can be done in consultation with the Food and Drug Administration.

13. The pharmaceutical company may use the CRT system 44 to design and carry out the Phase III study in accordance with Food and Drug Administration (FDA) requirements. The validation and replication of the Phase III results suggests an application to the marketplace gain new patients who have a high probability of favorable response to the compound.

14. The CRT system 44 may enable the pharmaceutical company to include the results of the Phase III pharmacogenomic study in the company's New Drug Submission to the FDA.

15. Other data from exploratory gene variants in the Phase II study may suggest the value of some but not all gene variants. The company may decide to utilize this information as part of its preclinical drug discovery program.

Example #2

A pharmaceutical company may pursue a discovery program which focuses on the discovery of small molecules which can be used to improve cognitive function in patients with Mild Cognitive Impairment (MCI, a precursor to Alzheimer's disease) and to alter the course and severity of Alzheimer's Disease itself. The CRT system 44 of the invention can be used to assist in such discovery.

It has now been established with good medical confidence that a variant of the Apolipoprotein Gene, APOE E4 allele, results in dose-dependent (homozygosity > heterozygosity) increased risk for Alzheimer's disease including early age of onset and diminished response to currently available therapeutic agents. Nevertheless, this variant is not believed to reflect the core etiology for Alzheimer's disease and many patients develop both MCI and Alzheimer's disease who do not have this allele. For this reason, the pharmaceutical company may wish to bring new chemical entities into clinical trials. The pharmaceutical company may also wish to examine the relationship between gene variants which code for enzymes (e.g. beta secretase) and proteins which are intrinsically involved in the Alzheimer's pathological processes.

1. The pharmaceutical company may use the CRT system 44 to conceptualize and design a clinical trial for their new chemical entity which incorporates pharmacogenomic principles.

2. Because of the known effects of the APOE E4 allele the risk and course of MCI and Alzheimer's disease, the company may wish to carry out its Phase II trials using a clinical population of MCI patients and those with early Alzheimer's disease in which there is equal representation of patients with and without the APOE E4 allele. The company may utilize the

CRT system 44 of the invention to create a pool of pre-genotyped patients who are then offered the opportunity to participate in the trial.

3. The pharmaceutical company may use the CRT system 44 to learn through the system's genotypic database 52 and retrieve additional information, for example, four other gene variants which have been reported to be related to the pathology of Alzheimer's disease.

4. As part of its exploratory Phase II study, the company may request patients to be pregenotyped for these four gene variants in addition to APOE E4 and decide to include four additional gene variants suggested by the company's scientists on the basis of pharmaceutical company's propriety discovery program.

5. The pharmaceutical company may utilize the CRT system 44 to provide information related to candidate genes, pharmacogenomic clinical trials design; it's selection of pre-genotyped patients is organized by the system. The CRT system 44 provides for the capacity to enter proprietary genetic information and to be used in the study.

6. In one scenario, the results of the phase II study may, for example, reveal a marked effect of APOB E4 status (negatively effecting treatment response) but also suggest, for example, two new gene variant predictors of favorable response.

7. Based upon these results, the pharmaceutical company may then design a Phase IIb study in which it intends to extend its early observations regarding new gene candidates, focusing on patients with MCI and Alzheimer's disease who do not have the APOE E4 allele. The CRT system 44 may further enable the pharmaceutical company to design the phase IIb study in accordance with FDA requirements.

8. The pharmaceutical company may then utilize the CRT system 44 to continue its development of this promising new chemical entity.

Example #3

A pharmaceutical company may be interested in pursuing treatment strategies for AIDS which will enhance current treatments aimed at delaying the onset of AIDS after an individual has positive immunoreactivity for the HIV virus. The CRT system 44 of the invention can be used to assist in such efforts.

Despite the fact that AIDS is an infectious disease caused by a retrovirus, genetic host factors, similar to other infectious diseases, can greatly influence the clinical course of the disorder. Specifically, genetic variants of several chemokine receptors which effect the function of an individual's immune system appear to delay the onset of AIDS following exposure to the HIV virus as reflected by positive immunoreactivity. For this reason, it is critical that the company control for known genetic causes of delayed AIDS onset in its treatment population in order to accurately determine its drug's therapeutic effects.

1. The pharmaceutical company may employ the CRT system 44 to determine known genetic variants known to alter the time course for the onset of AIDS.

2. The pharmaceutical company may wish to enter subjects with positive immunoreactivity to the HIV virus and be assured that genetic host factors are equally represented in all arms of the study.

3. The pharmaceutical company may employ the CRT system 44 to develop a pool of candidate patients who are also pre-genotyped for chemokine receptor gene variants. As patients continue to enter the study over the expected three years duration, the CRT system 44 may monitor the balance of genotypes in treatment arms.

4. The results of the study after, for example, three years may reveal that combination of "protective" genotypes and the experimental therapeutic agent result in <1%

conversion to AIDS in comparison with a 25% conversion rate for patients without "protective" genotypes and who received the active experimental drug. Because of ethical considerations, placebo will not be included in the study but placebo data, transposed from large public health AIDS databases, may reveal >50% conversion rate for genetically cross-sectional analysis of positive immunreactive patients.

5. The pharmacogenomic approach which is enabled by the CRT system 44 may result in demonstration of synergistic effect of the experimental agent with host genetic factors and support the agent's superiority to the natural disease progression.

Example #4

A pharmaceutical company may wishe to bring a lead compound targeted as an antipsychotic into clinical trials. The system of the invention can be used to assist in such efforts.

In this example, the compound has already passed through Phase I trials and showed no limiting adverse events in normal controls. Moreover, Phase II trial data suggest that the drug has an impressive antipsychotic profile. It is noted, however, that a potential market limiting adverse event was observed in some, but not all subjects: weight gain > 4 kg over 6 weeks of drug administration. It is known that weight gain limits market acceptance and further may predispose diabetes. For this reason, the company wishes to carry out its Phase III trial in a fashion that minimizes weight gain without interference with the signal of the drug's effectiveness. Once efficacy is established, the company wishes to demonstrate that those subjects at risk for weight gain may benefit from co-administration of a second, commercially available agent which otherwise may be of no benefit to the indication (e.g., HI antagonist).

The pharmaceutical company may utilize the pharmacogenomic therapy recommendation system 300 of the invention in the following fashion:

1. The pharmacogenomic therapy recommendation system 300 first provides the company with information regarding drug-induced weight gain including known genetic variants which are thought to represent risk factors.
2. The pharmacogenomic therapy recommendation system 300 also provides information regarding specific genetic associations between weight gain in relation to the category and chemical class of the company's drug undergoing clinical trial.
3. It is learned using the pharmacogenomic therapy recommendation system 300 that variants of the 5HTzc receptors are associated with weight gain in patients. This becomes relevant because the drug in clinical trial has, as part of its mechanism of action, antagonist properties to the 5HTzc receptor.
4. It is also learned using the pharmacogenomic therapy recommendation system 300 that there are 200 low frequency SNP's which have been found in individual patients to be associated with excessive weight gain to a number of different drugs.
5. The company desires to target the major emphasis of its Phase III clinical population to exclude patients with risk of excessive weight gain, so as establish clear effectiveness as an antipsychotic without limiting market possibilities.
6. As a secondary goal, the company whishes to conduct a preliminary, controlled trial in patients at risk for excessive weight gain which includes co-administration of an H1 antagonist to minimize the anticipated adverse event.
7. The strategy proves to be successful in that the clinical trial excluding patients at risk for excessive weight gain event achieves the desired goal: antipsychotic efficacy in demonstrated with modest group weight gain data (4.5kg per patient).

8. The company continues with the clinical trial designed specifically for at risk patients and finds that co-administration restricts weight gain without interfering with antipsychotic efficacy.
9. The NDA application includes genotypic patient information which provides a risk assessment for excessive weight gains dependent upon genotype.
10. The company utilizes the pharmacogenomic therapy recommendation system 300 to develop diagnostic DNA "chip" enabling all patients to be candidates for treatment with their drug and allows for identification of those at risk for weight gain, enabling a strategy of co-administration of a secondary agent.

Example #5

In this prophetic example, a company has successfully marketed a broad spectrum antibiotic. It has become apparent that an unexpected and rare adverse event has emerged which has potentially serious implications for continued market success. This event is the emergence of psychosis and/or other serious behavioral disturbances during treatment >14 days. The drug has now received increasing attention and unanticipated application as a prophylactic agent for the fatal disorder, Anthrax. The drug is administered to many more patients and for a longer period of time than expected. The behavioral adverse event takes on greater significance as the drug is now widely used for extensive treatment periods in individuals exposed to this toxic agent and enhances fear related to potential anthrax exposure. Because other antibiotics without this adverse event may also be effective in the treatment of Anthrax, the company wishes to identify patients at risk for behavioral adverse events.

The pharmaceutical company may use the pharmacogenomic therapy recommendation system 300 and CTR system 44 to accomplish their goal of establishing genetic risk factors for the relatively rare but limiting adverse event of toxic behavioral disturbance.

1. The company first gains access to the pharmacogenomic therapy recommendation system 300 to learn about known genetic factors which might predispose to psychotic reactions.
2. Some of the information received relates to targets, genes and gene variants known to be associated with the neurobiology of psychosis in general.
3. The CTR system 44 confirms that drug-induced behavioral disturbances occur with treatment with other drug categories, although the known mechanisms (e.g., inhibition of dopamine beta-hydroxylase) do not appear to be related to pharmacological effects of the company's antibiotic.
4. The company decides to utilize the CTR system 44 in two ways. First, it will genotype patients who have experienced psychosis during treatment with the antibiotic for targets thought to represent neuronal pathways for psychosis. These include, but are not limited to known variants of genes coding for dopamine receptors (2,3,4) and the dopamine transporter gene.
5. The company also decides to access from the CTR system 44 (and add new patients to the database) patients who have experienced drug-induced psychosis to the antibiotic. The company then proceeds with an intensive deep sequencing program designed to identify heretofore unknown low frequency SNP variants that are associated with drug-induced psychosis.
6. Utilizing the CTR system 44, the company establishes that all individuals who have experienced drug-induced psychosis to the antibiotic show a far higher frequency, of

functional gene variants of the dopamine transporter than are found in the general population. This leads to the hypothesis that gene variants of the dopamine transporter are a major determinant in providing risk for antibiotic-induced psychosis.

7. The company begins Phase IV trials to prospectively test the hypothesis of the role of the dopamine transporter variant in the induced psychosis produced by its drug.
8. The company further discovers through its sequencing program which utilizes the CTR system 44 database that there is a pattern of low frequency SNP's which also are related to behavioral disturbances produced by the antibiotic.
9. The company initiates the development of a proprietary DNA chip intended to identify patients at risk for behavioral disturbances associated with treatment with its proprietary antibiotic.

Example #6

In this prophetic example, a company has developed a drug which has received approval by the FDA as branded prescription product. During the conduct of Phase III trials, it was noted that mean increases in the QTc EKG interval occur but it was agreed by the FDA, including its advisory panel, to be a drug-induced physiological effect that represented an acceptable risk, comparable to other drugs of varying therapeutic categories which already had received approval. The company's drug proved to be very successful in the clinical market. However, unexpectedly several serious cardiac arrythmias were reported in Phase IV monitoring by FDA - two cases resulted in sudden death. Because of the severity of these adverse events, immediate scrutiny into the drug's safety was launched by both regulatory and company personnel.

The company utilized the CTR system 44 in order to establish a better understanding of genetic risk factors associated with drug-induced arrhythmias in order to "rescue" an otherwise therapeutically sound and commercially successful product.

1. The company's utilization of the CTR system 44 rapidly identified risk factors discovered through studies of Sudden Infant Death Syndrome which may represent occult risk to drug-induced cardiac arrhythmias and other cardiac conduction defects in adults.
2. The CTR system 44 identified variants of the SCN5A sodium channel gene to be associated with cardiac arrhythmias in 30% of its database of drug-induced cardiac arrhythmias in adults. Moreover, the CTR system 44 provided a sample of 500 patients who had documented drug-induced prolongation of the QTc interval and thus were at high risk for cardiac arrhythmia and sudden death.
3. The drug-induced QTc population patient group was suited for prospective investigation into genetic cardiac risk factors which could have implications to the company's drug.
4. The company rapidly performed genotyping for the SCN5A gene on subjects who had experienced drug-induced arrhythmia in Phase IV monitoring. The company found a frequency of 40% in this group, far above the general population frequency of <2%. This information provided a first step to developing a strategy to "rescue" their drug.
5. The CTR system 44's biological database was utilized by the company to extend its research. In addition to the SCNSA, the company discovered through deep sequencing programs that SNP's in other ion channel related genes were found in the drug-induced QTc prolongation population.

6. The company returned to the original group of patients who had experienced cardiac arrhythmias to its drug and examined the frequencies of the newly discovered SNP's. Several but not all of the SNP's were found in the samples.
7. The company initiates a Phase IV trial utilizing these pharmacogenetics findings to validate its hypothesis that these genes represent risk factors for the adverse event of their drug.
8. The FDA's post marketing surveillance policy supports the company's strategy regarding identification of genetic risk factors.
9. The company requests the services of a DNA array chip manufacturer to design a "chip" expressively for the purpose of identifying patients at risk for drug-induced cardiac arrhythmia produced by their drug.
10. The company enters into discussion with the FDA regarding the approval process for this DNA chip as a diagnostic entity.

We Claim:

1. A pharmacogenomic system comprising:

a genotype database (GDB), the GDB comprising genetic information for a plurality of patients;

a clinical database (CDB), the CDB comprising clinical phenotypic information for a plurality of patients;

a clinical trial requirement database (CRDB), the CRDB comprising information on clinical trial requirements for at least one phase of a clinical trial;

an association module connected to GDB, and CDB that is adapted to determine an association between the genetic information and the clinical phenotypic information for a plurality of patients; and

a recommendation module connected to GDB, CDB, and CRDB that is adapted to provide clinical trial recommendations utilizing the genetic information, the clinical phenotypic information, the clinical trial requirement information and the determined association between the clinical information and the genetic information.

2. The system of claim 1 further comprising a selection module that is connected to the system for selecting one or more patients based on the genetic information, wherein the selection is performed using plurality of statistical methods.

3. The system of claim 1, wherein the genetic information correspond to one or more variation in candidate genes.

4. The system of claim 1, wherein the genetic information correspond to plurality of Single Nucleotide Polymorphisms.

5. The system of claim 1, wherein the clinical trial requirement information correspond to one or more protocols for Phase I of a clinical trial.

6. The system of claim 1, wherein the clinical trial requirement information correspond to one or more inclusion/exclusion criteria for clinical trials.

7. The system of claim 1, wherein the association is determined my one or more of pre-determined statistical methods.

8. The system of claim 1, wherein the recommendation modules perform optimization of clinical trial parameters for providing clinical trial recommendations.

9. A pharmacogenomic system comprising:

a genotype database (GDB), the GDB comprising genetic information for a plurality of patients;

a clinical database (CDB), the CDB comprising clinical phenotypic information for a plurality of patients;

a clinical trial requirement database (CRDB), the CRDB comprising information on clinical trial requirements for at least one phase of a clinical trial;

association means connected to GDB, and CDB for determining an association between the genetic information and the clinical phenotypic information for a plurality of patients; and

recommendation means connected to GDB, CDB, and CRDB, for providing clinical trial recommendations utilizing the genetic information, the clinical phenotypic information, the clinical trial requirement information and the determined association between the clinical information and the genetic information.

10. The system of claim 9 further comprising selection means for selecting one or more patients based on the genetic information, wherein the selection is performed using plurality of statistical methods.

11. The system of claim 9, wherein the genetic information correspond to one or more variation in candidate genes.

12. The system of claim 9, wherein the genetic information correspond to plurality of Single Nucleotide Polymorphisms.

13. The system of claim 9, wherein the clinical trial requirement information correspond to one or more protocols for Phase I of a clinical trial.

14. The system of claim 9, wherein the clinical trial requirement information correspond to one or more inclusion/exclusion criteria for clinical trials.

15. The system of claim 9, wherein the association is determined my one or more of pre-determined statistical methods.

16. The system of claim 9, wherein the recommendation means perform optimization of clinical trial parameters for providing clinical trial recommendations.

17. A pharmacogenomic method comprising the steps of:

enabling a user to access a genotype database (GDB), the GDB comprising genetic information for a plurality of patients;

enabling a user to access a clinical database (CDB), the CDB comprising clinical phenotypic information for a plurality of patients;

enabling a user to access a clinical trial requirement database (CRDB), the CRDB comprising information on clinical trial requirements for at least one phase of a clinical trial;

enabling a user to determine an association between the genetic information and the clinical phenotypic information for a plurality of patients; and

enabling a user to cause the system to provide clinical trial recommendations utilizing the genetic information, the clinical phenotypic information, the clinical trial requirement information and the determined association between the clinical information and the genetic information.

18. The method of claim 17 further comprising a step of enabling a user to select one or more patients based on the genetic information, wherein the selection is performed using plurality of statistical methods.

19. The method of claim 17, wherein the genetic information correspond to one or more variation in candidate genes.

20. The method of claim 17, wherein the genetic information correspond to plurality of Single Nucleotide Polymorphisms.

21. The method of claim 17, wherein the clinical trial requirement information correspond to one or more protocols for Phase I of a clinical trial.

22. The method of claim 17, wherein the clinical trial requirement information correspond to one or more inclusion/exclusion criteria for clinical trials.

23. The method of claim 17, wherein the association is determined my one or more of pre-determined statistical methods.

24. The method of claim 17, wherein method performs optimization of clinical trial parameters for providing clinical trial recommendations.

25. A processor readable pharmacogenomic medium comprising:

a first processor readable program code for enabling a user to access a genotype database (GDB), the GDB comprising genetic information for a plurality of patients;

a second processor readable program code for enabling a user to access a clinical database (CDB), the CDB comprising clinical phenotypic information for a plurality of patients;

a third processor readable program code for enabling a user to access a clinical trial requirement database (CRDB), the CRDB comprising information on clinical trial requirements for at least one phase of a clinical trial;

a fourth processor readable program code for enabling a user to determine an association between the genetic information and the clinical phenotypic information for a plurality of patients; and

a fifth processor readable program code for enabling a user to cause the system to provide clinical trial recommendations utilizing the genetic information, the clinical phenotypic information, the clinical trial requirement information and the determined association between the clinical information and the genetic information.

26. A pharmacogenomic system comprising:

a genotype database (GDB), the GDB comprising genetic information for a plurality of patients;

an adverse drug event database (AEDB), the AEDB comprising adverse drug event phenotypic information of for a plurality of patients;

an association module connected to GDB and AEDB and adapted to enable a user to determine an association between the genetic information and the adverse drug event phenotypic information for a plurality of patients;

a risk prediction module that enables a user to predict a risk for adverse drug events for a plurality of patients, wherein the risk prediction module utilizes the determined association between the genetic information and the adverse drug event phenotypic information;

a validation module that enables a user to validate the predicted risk for adverse drug events for a plurality of patients; and

a recommendation module that enables a user to recommend prescription utilizing the validated information for risk for adverse drug events for a plurality of patients.

27. The system of claim 26 further comprising a selection module for selecting one or more patients based on the genetic information, wherein the selection is performed using plurality of statistical methods.

28. The system of claim 26, wherein the genetic information correspond to one or more variation in candidate genes.

29. The system of claim 26, wherein the genetic information correspond to plurality of Single Nucleotide Polymorphisms.

30. The system of claim 26, wherein adverse drug events correspond to multiple physiological systems with multiple clinical manifestations.

31. The system of claim 26, wherein the association is determined my one or more of pre-determined statistical methods.

32. The system of claim 26, wherein the validation is performed utilizing one or more pre-determined mathematical models.

33. A pharmacogenomic method comprising the steps of:

enabling a user to access a genotype database (GDB), the GDB comprising genetic information for a plurality of patients;

enabling a user to access a adverse drug event database (AEDB), the AEDB comprising adverse drug event phenotypic information of for a plurality of patients;

enabling a user to determine an association between the genetic information and the adverse drug event phenotypic information for a plurality of patients;

enabling a user to predict a risk for adverse drug events for a plurality of patients, wherein the risk prediction modules utilize the determined association between the genetic information and the adverse drug event phenotypic information;

enabling a user to validate the predicted risk for adverse drug events for a plurality of patients; and

enabling a user to recommend prescription utilizing the validated information for risk for adverse drug events for a plurality of patients.

34. The method of claim 33 further comprising the step of selecting one or more patients based on the genetic information, wherein the selection is performed using plurality of statistical methods.

35. The method of claim 33, wherein the genetic information correspond to one or more variation in candidate genes.

36. The method of claim 33, wherein the genetic information correspond to plurality of Single Nucleotide Polymorphisms.

37. The method of claim 33, wherein adverse drug events correspond to multiple physiological systems with multiple clinical manifestations.

38. The method of claim 33, wherein the association is determined my one or more of pre-determined statistical methods.

39. The method of claim 33, wherein the validation is performed utilizing one or more pre-determined mathematical models.

40. A processor readable pharmacogenomic medium comprising:

a first processor readable program code for enabling a user to access a genotype database (GDB), the GDB comprising genetic information for a plurality of patients;

a second processor readable program code for enabling a user to access a adverse drug event database (AEDB), the AEDB comprising adverse drug event phenotypic information of for a plurality of patients;

a third processor readable program code for enabling a user to determine an association between the genetic information and the adverse drug event phenotypic information for a plurality of patients;

a fourth processor readable program code for enabling a user to predict a risk for adverse drug events for a plurality of patients, wherein the risk prediction modules utilize the determined association between the genetic information and the adverse drug event phenotypic information;

a fifth processor readable program code for enabling a user to validate the predicted risk for adverse drug events for a plurality of patients; and

a sixth processor readable program code for enabling a user to recommend prescription utilizing the validated information for risk for adverse drug events for a plurality of patients.

41. A pharmacogenomic system for predicting a risk of adverse events to one or more drugs for a plurality of patients, the system comprising:

means for providing genetic information for a plurality of patients;

means for providing adverse drug event phenotypic information of for a plurality of patients;

means for enabling a user to determine an association between the genetic information and the adverse drug event phenotypic information for a plurality of patients;

risk prediction means that enable a user to predict a risk for adverse drug events for a plurality of patients, wherein the risk prediction modules

utilize the determined association between the genetic information and the adverse drug event phenotypic information;

validation means that enable a user to validate the predicted risk for adverse drug events for a plurality of patients; and

recommendation means that enable a user to recommend prescription utilizing the validated information for risk for adverse drug events for a plurality of patients.

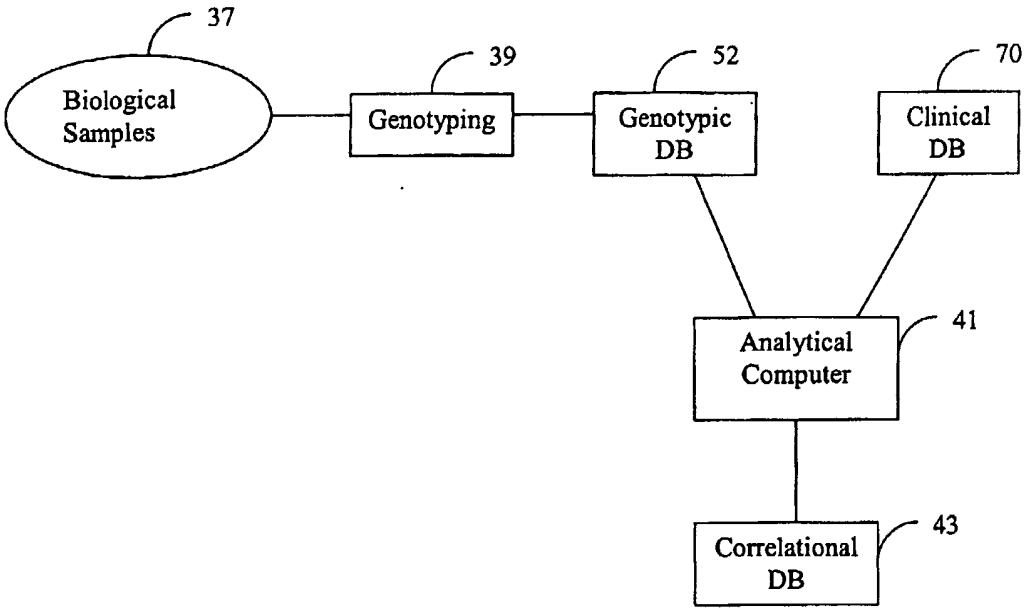
37
Biological Samples
39
Genotyping
52
Genotypic DB
70
Clinical DB
41
Analytical Computer
43
Correlational DB

Fig. 1

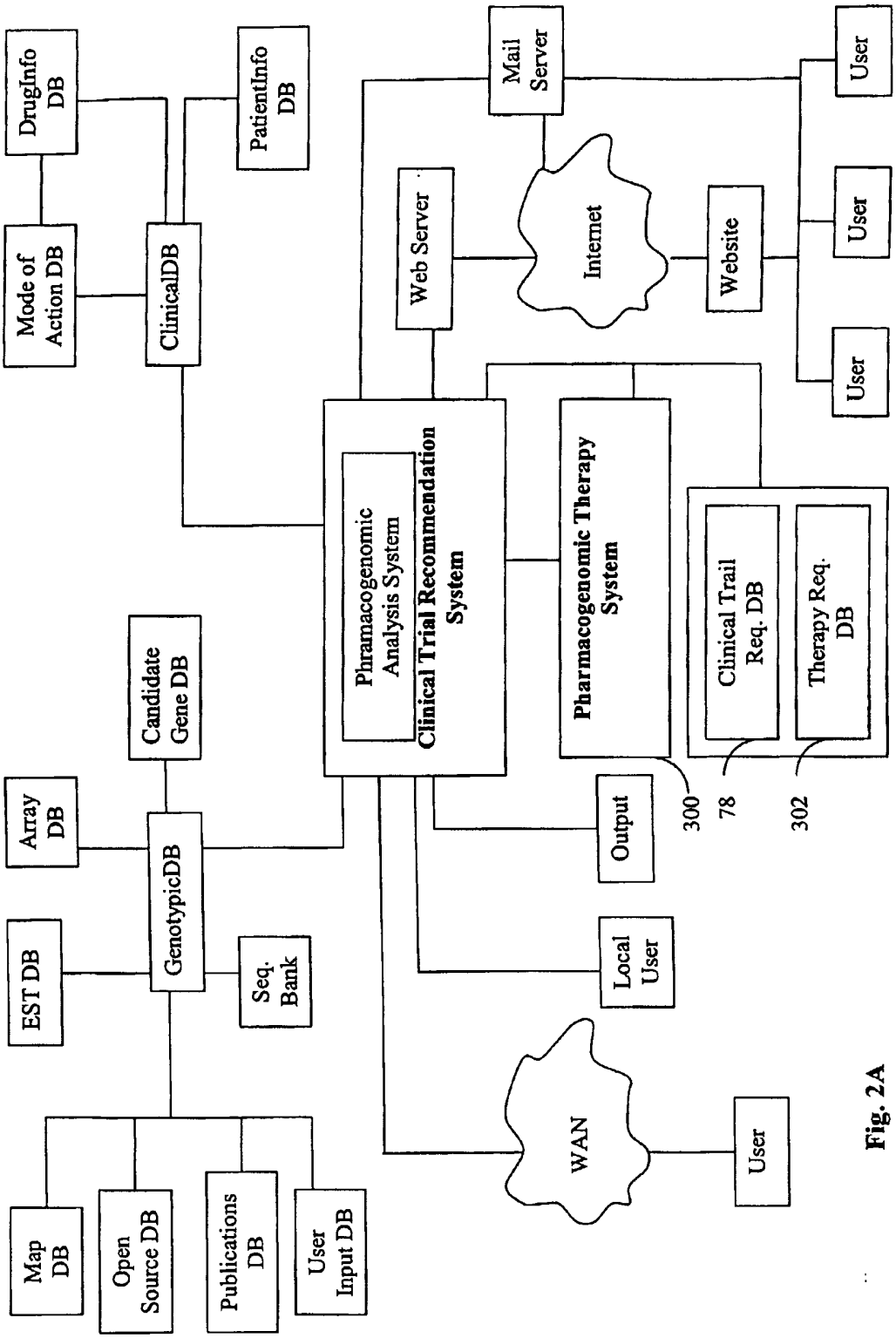
DrugInfo DB
PatientInfo DB
Mail Server
User
Mode of Action DB
ClinicalDB
Web Server
Internet
Website
User
User
Phramacogenomic Analysis System
Clinical Trial Recommendation System
Pharmacogenomic Therapy System
Clinical Trail Req. DB
Therapy Req. DB
Candidate Gene DB
Array DB
GenotypicDB
Output
300
78
302
EST DB
Seq. Bank
Local User
WAN
User
Map DB
Open Source DB
Publications DB
User Input DB

Fig. 2A

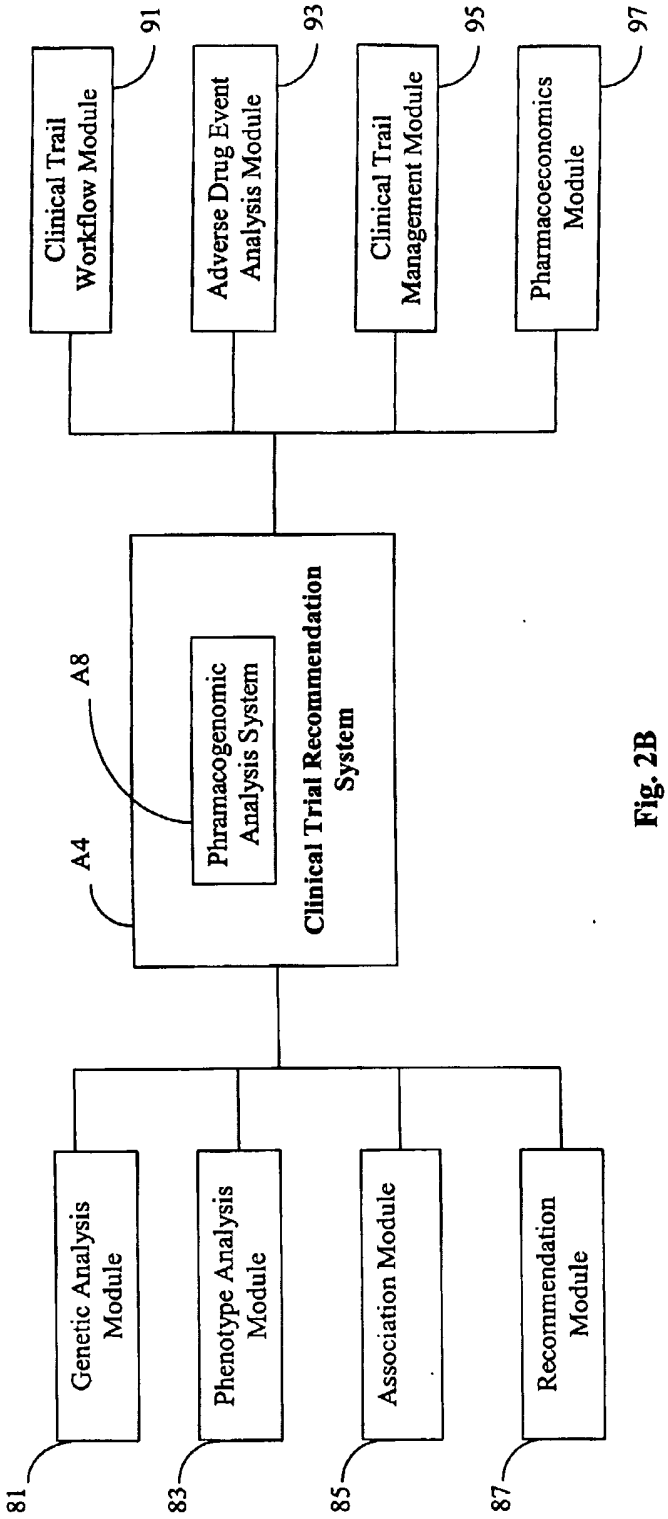
Clinical Trail Workflow Module
91
Adverse Drug Event Analysis Module
93
Clinical Trail Management Module
95
Pharmacoeconomics Module
97
A8
Phramacogenomic Analysis System
A4
Clinical Trial Recommendation System
Genetic Analysis Module
81
Phenotype Analysis Module
83
Association Module
85
Recommendation Module
87

Fig. 2B

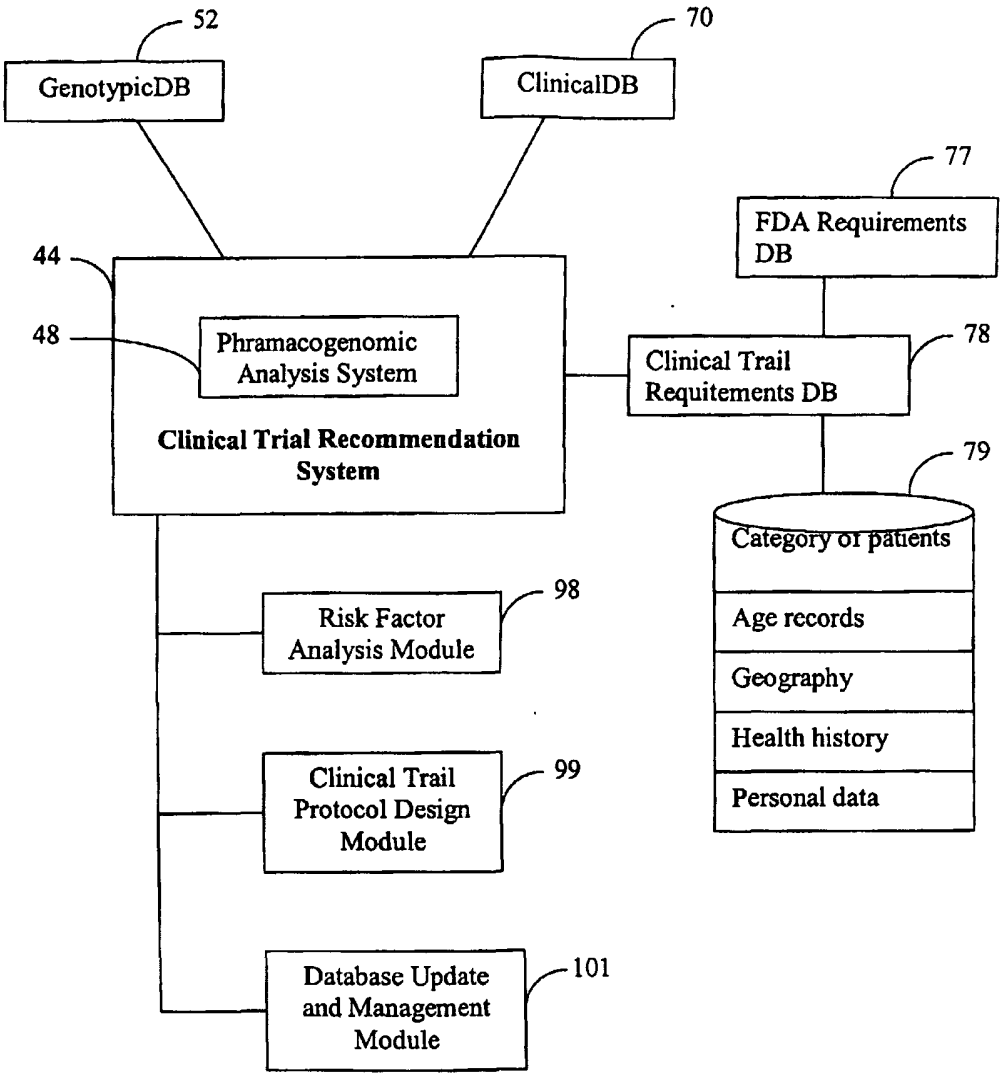
52
GenotypicDB
70
ClinicalDB
77
FDA Requirements DB
44
48
Phramacogenomic Analysis System
Clinical Trial Recommendation System
Clinical Trail Requitements DB
78
79
Category of patients
Age records
Geography
Health history
Personal data
98
Risk Factor Analysis Module
99
Clinical Trail Protocol Design Module
101
Database Update and Management Module

Fig. 2C

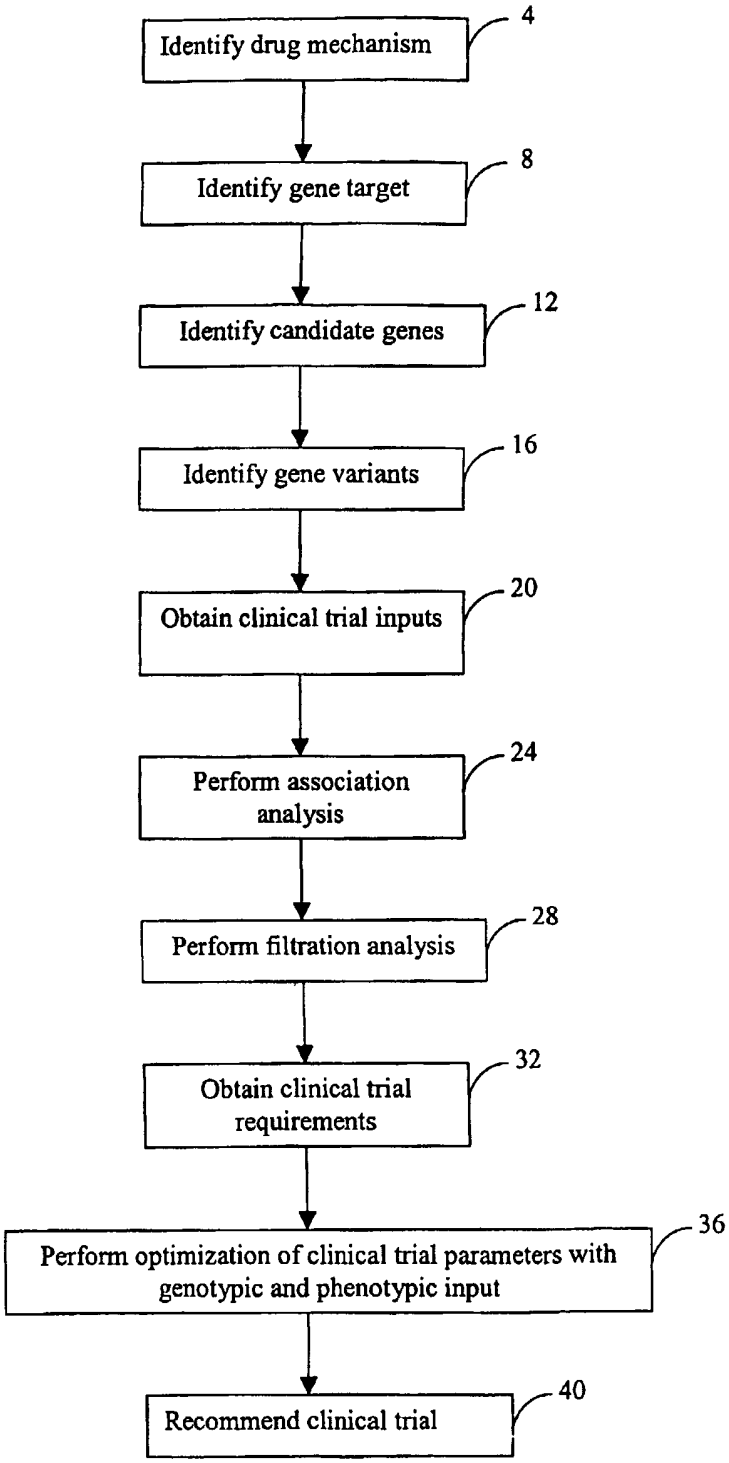
Identify drug mechanism
4
Identify gene target
8
Identify candidate genes
12
Identify gene variants
16
Obtain clinical trial inputs
20
Perform association analysis
24
Perform filtration analysis
28
Obtain clinical trial requirements
32
Perform optimization of clinical trial parameters with genotypic and phenotypic input
36
Recommend clinical trial
40

Fig. 3A

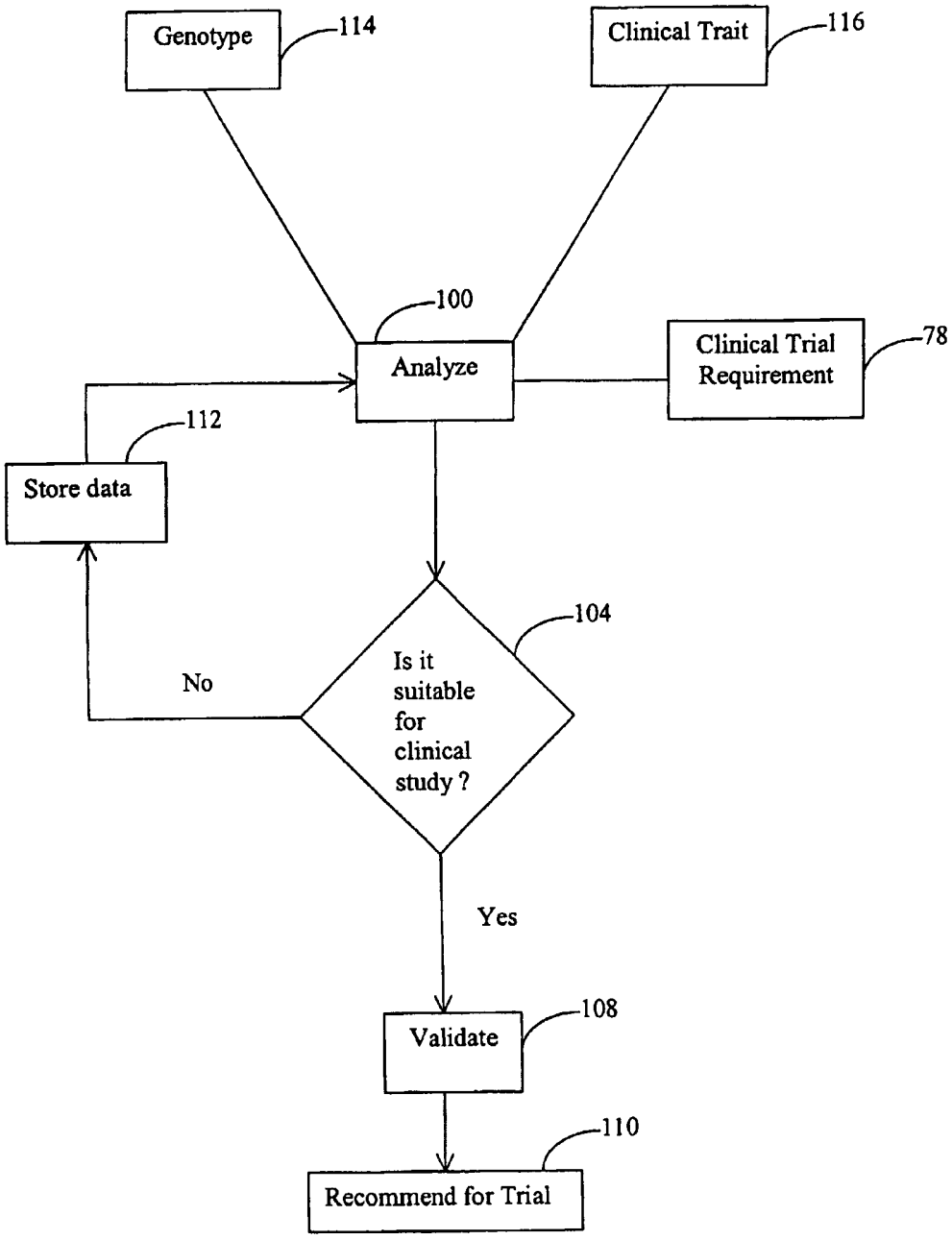
Genotype
114
Clinical Trait
116
100
Analyze
Clinical Trial Requirement
78
112
Store data
104
Is it suitable for clinical study ?
No
Yes
108
Validate
110
Recommend for Trial

Fig. 3B

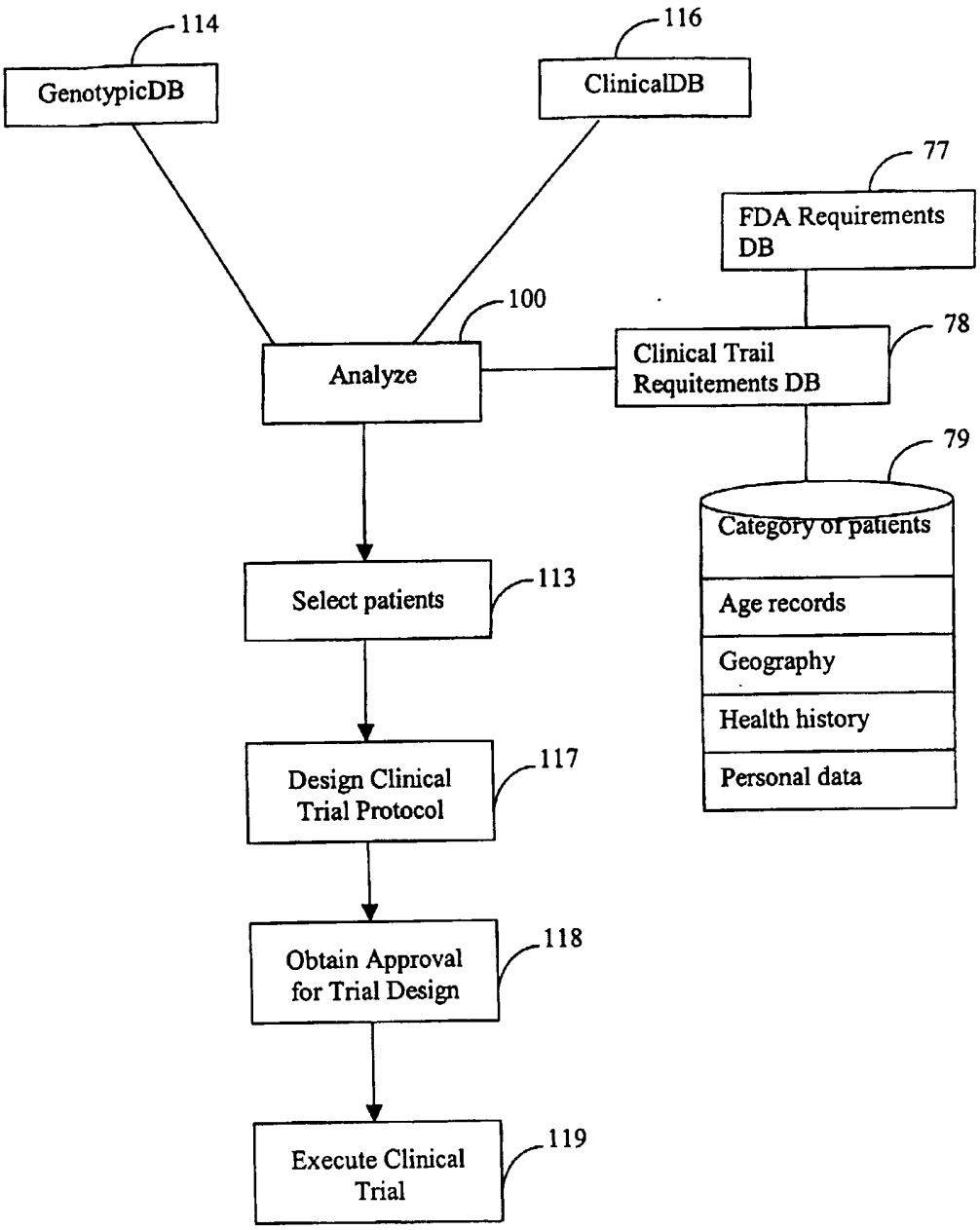
114
GenotypicDB
116
ClinicalDB
77
FDA Requirements DB
100
Analyze
78
Clinical Trail Requitements DB
79
Category of patients
Age records
Geography
Health history
Personal data
Select patients
113
Design Clinical Trial Protocol
117
Obtain Approval for Trial Design
118
Execute Clinical Trial
119

Fig. 3C

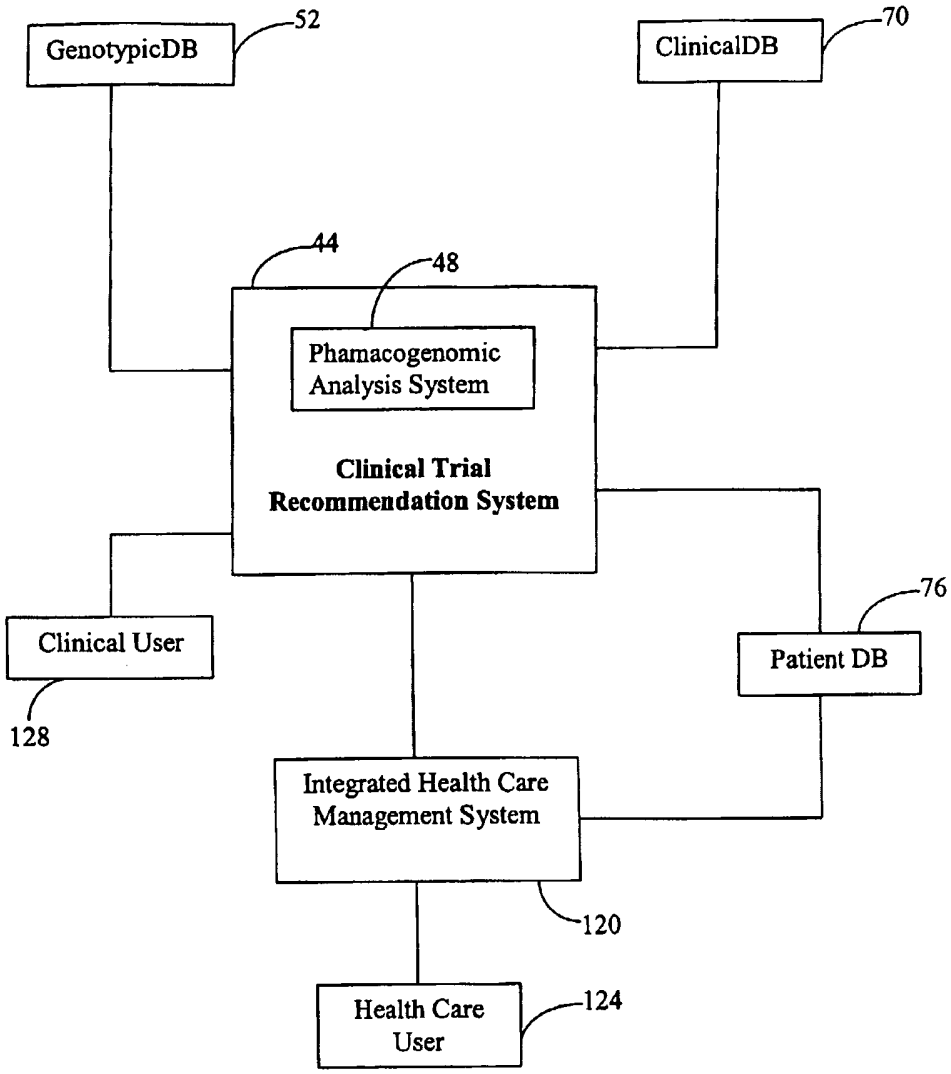
GenotypicDB
52
ClinicalDB
70
44
48
Phamacogenomic Analysis System
Clinical Trial Recommendation System
Clinical User
128
76
Patient DB
Integrated Health Care Management System
120
Health Care User
124

Fig. 4A

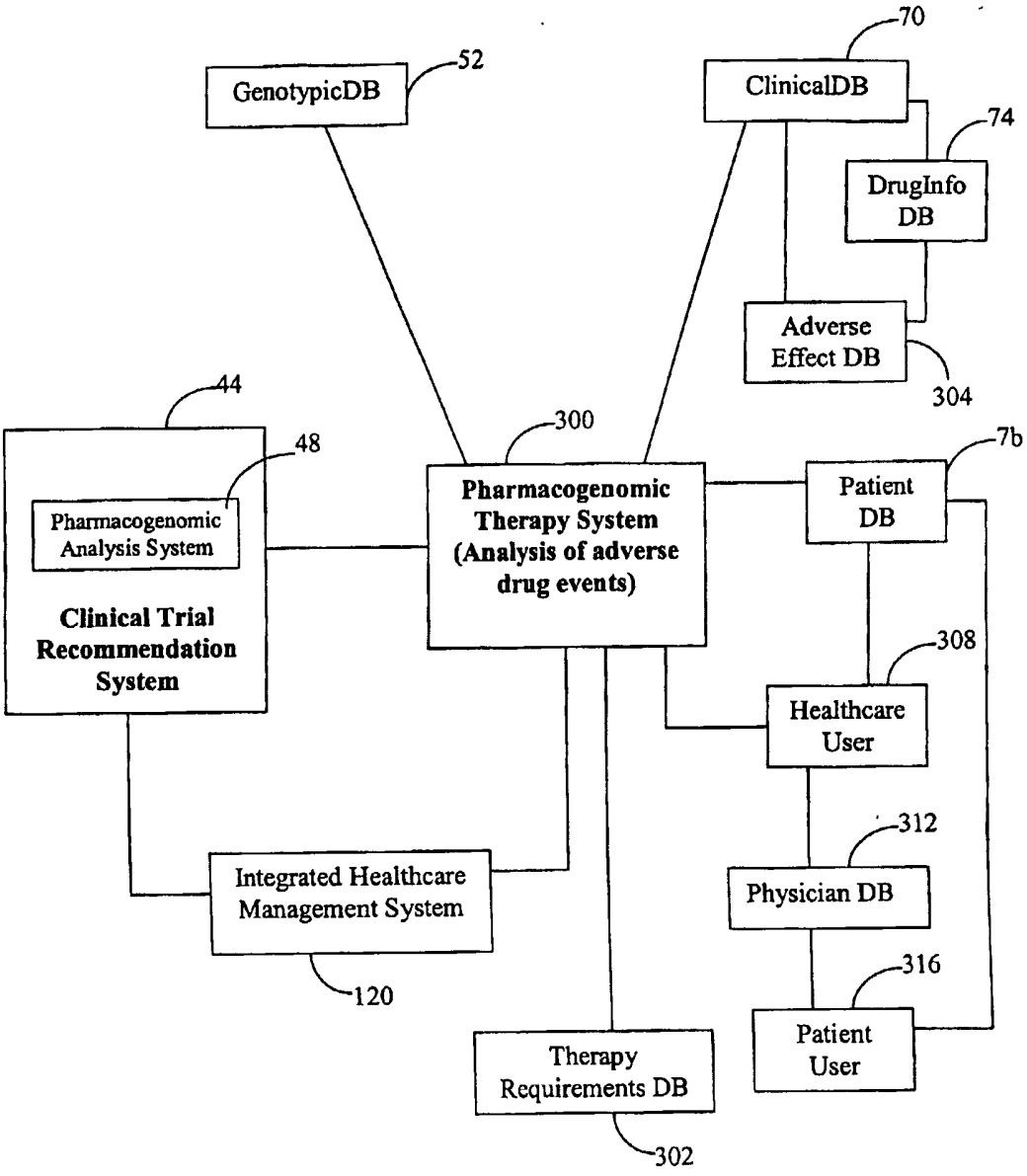
GenotypicDB
52
ClinicalDB
70
74
DrugInfo DB
Adverse Effect DB
304
44
48
300
7b
Pharmacogenomic Analysis System
Clinical Trial Recommendation System
Pharmacogenomic Therapy System (Analysis of adverse drug events)
Patient DB
308
Healthcare User
312
Integrated Healthcare Management System
120
Physician DB
316
Patient User
Therapy Requirements DB
302

Fig. 4B

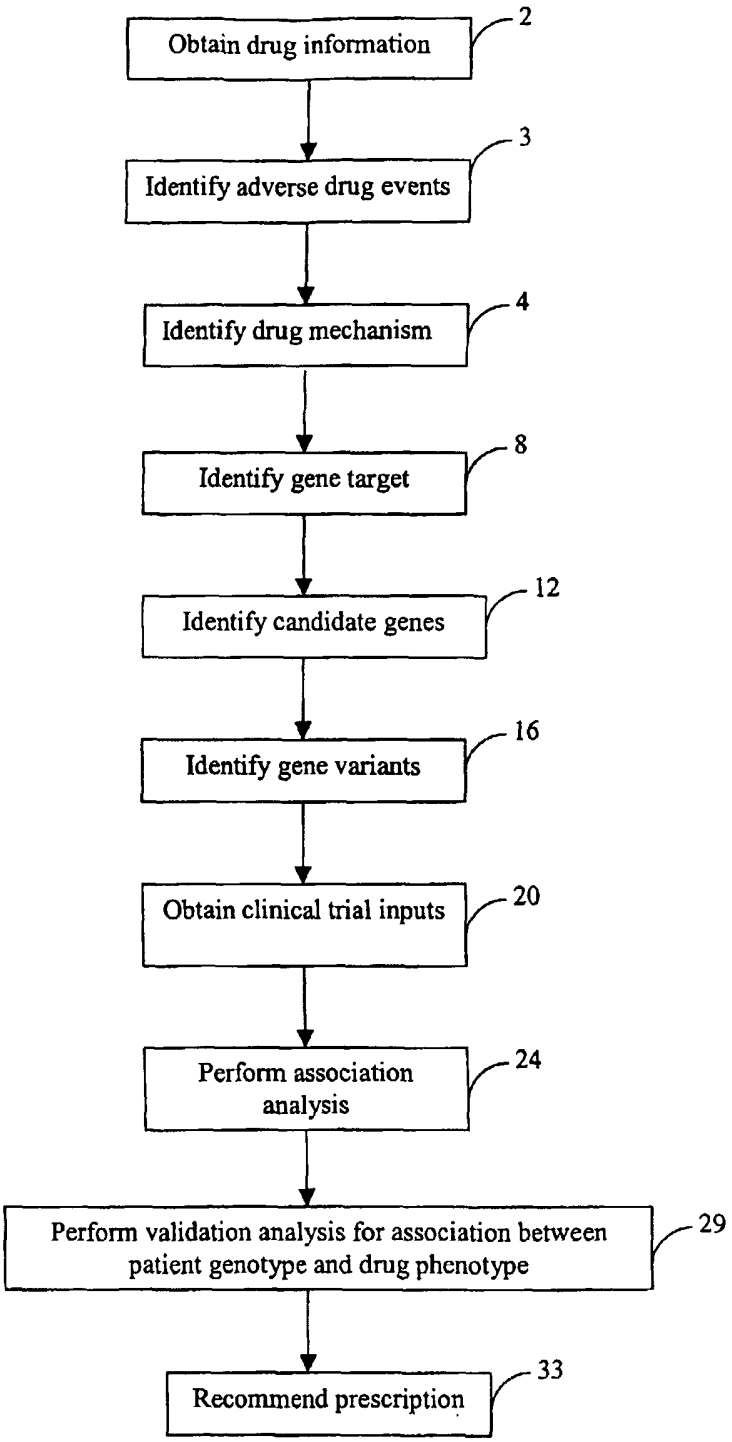
Obtain drug information
2
Identify adverse drug events
3
Identify drug mechanism
4
Identify gene target
8
Identify candidate genes
12
Identify gene variants
16
Obtain clinical trial inputs
20
Perform association analysis
24
Perform validation analysis for association between patient genotype and drug phenotype
29
Recommend prescription
33

Fig. 4C

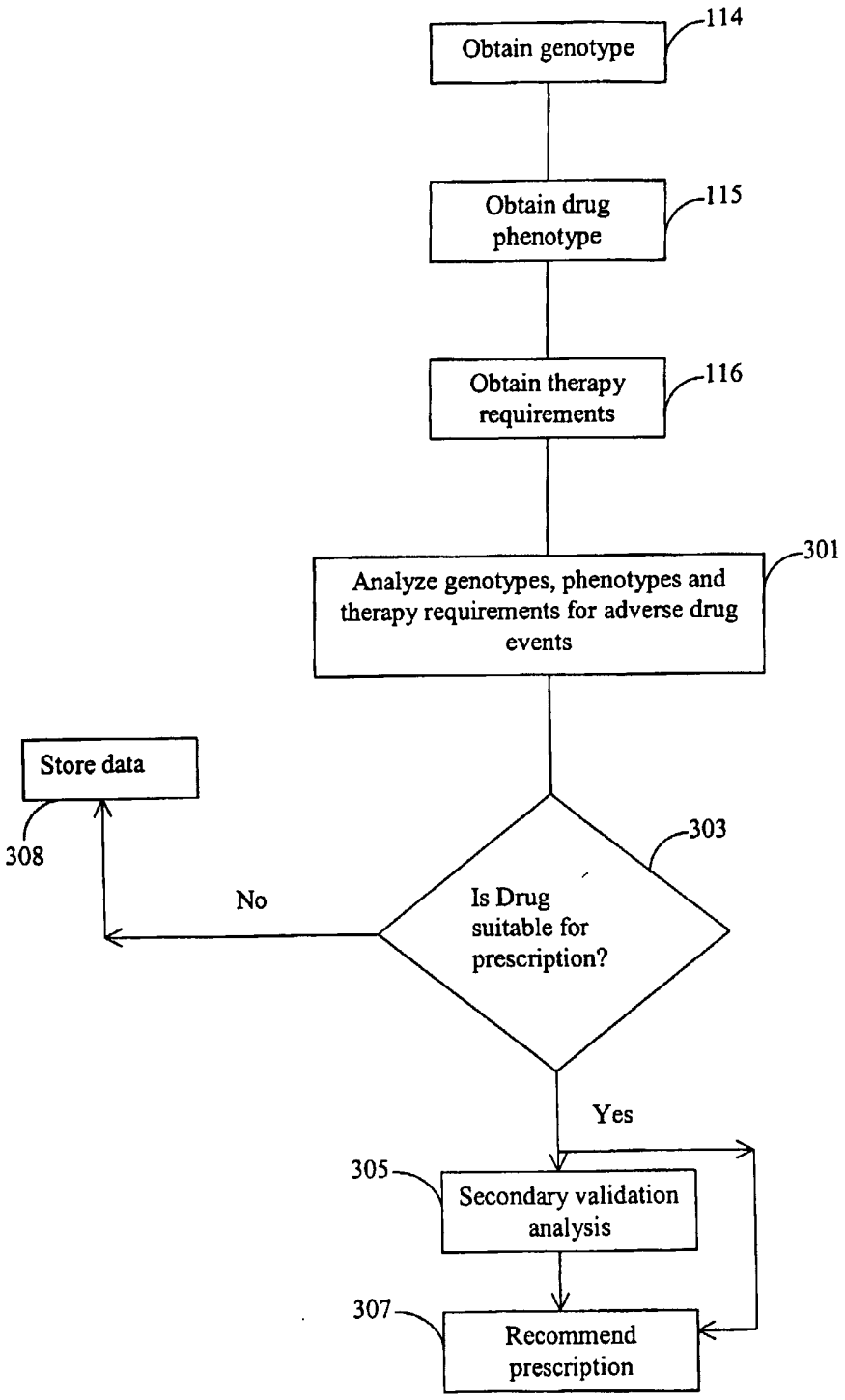
Obtain genotype
114
Obtain drug phenotype
115
Obtain therapy requirements
116
Analyze genotypes, phenotypes and therapy requirements for adverse drug events
301
Store data
308
No
Is Drug suitable for prescription?
303
Yes
305
Secondary validation analysis
307
Recommend prescription

Fig. 4D

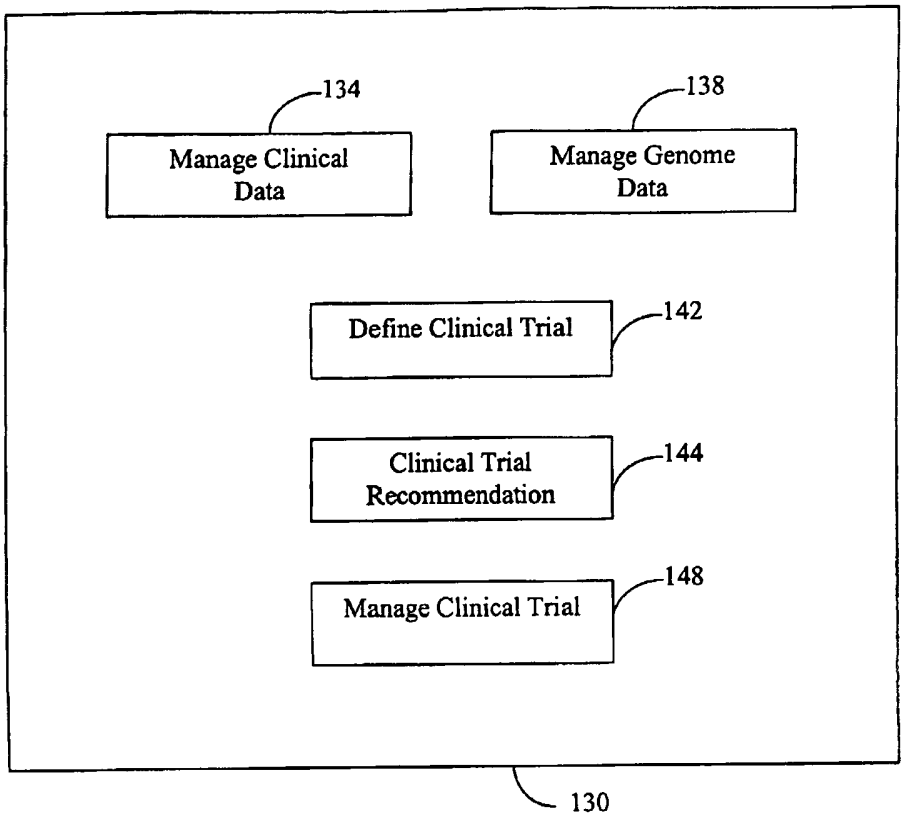
134
Manage Clinical Data
138
Manage Genome Data
Define Clinical Trial
142
Clinical Trial Recommendation
144
Manage Clinical Trial
148
130

Fig. 5A

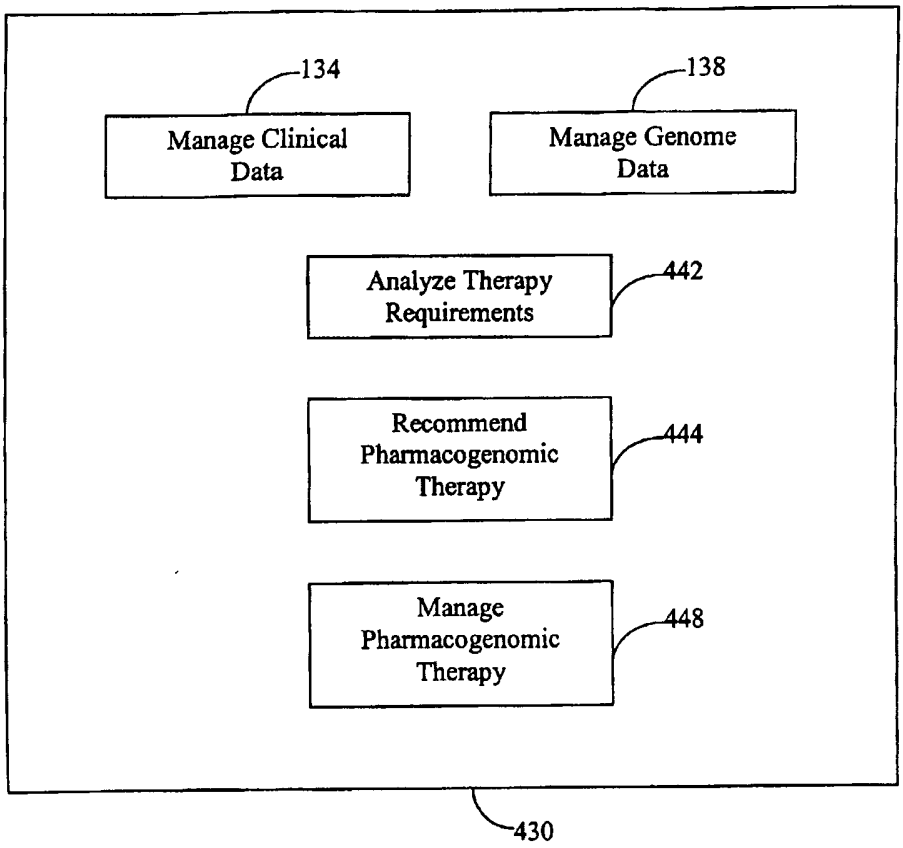
134
Manage Clinical Data
138
Manage Genome Data
Analyze Therapy Requirements
442
Recommend Pharmacogenomic Therapy
444
Manage Pharmacogenomic Therapy
448
430

Fig. 5B

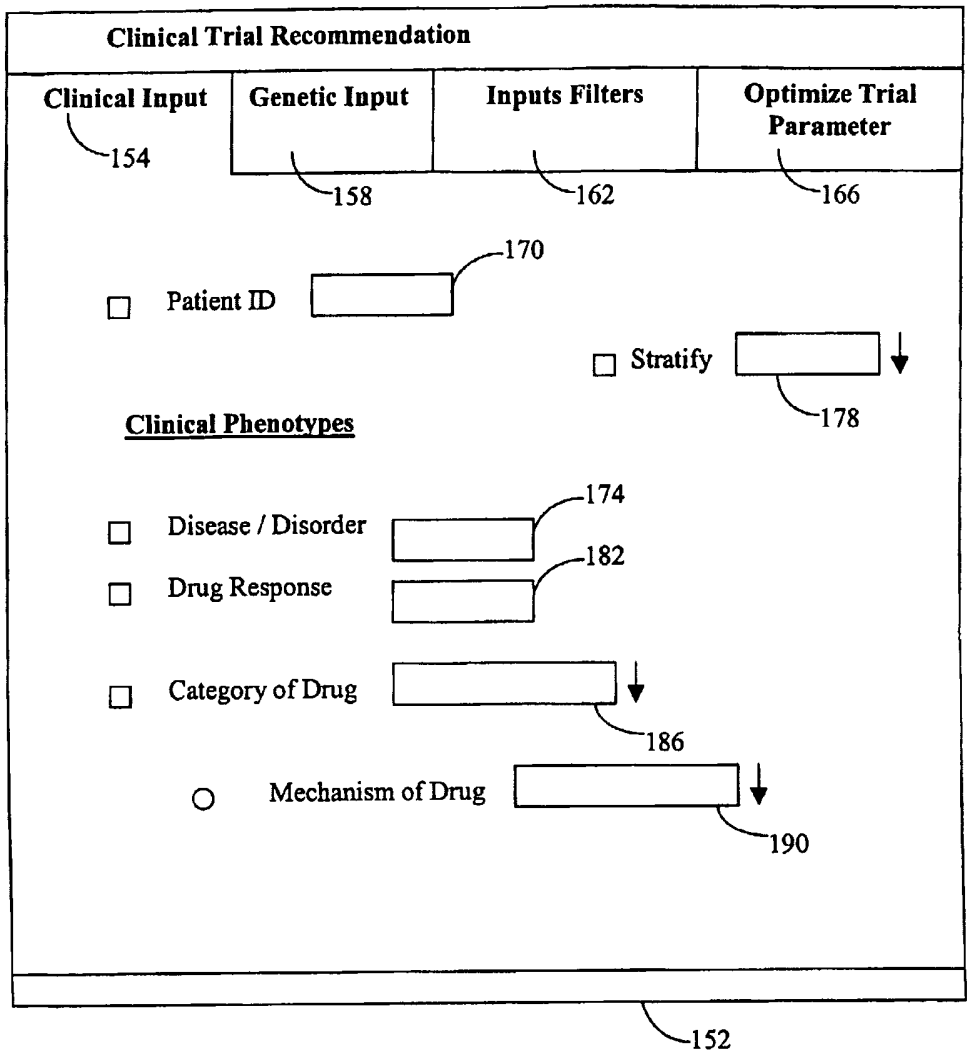
Clinical Trial Recommendation
Clinical Input
Genetic Input
Inputs Filters
Optimize Trial Parameter
154
158
162
166
170
Patient ID
Stratify
178
Clinical Phenotypes
174
Disease / Disorder
182
Drug Response
Category of Drug
186
Mechanism of Drug
190
152

Fig. 5C

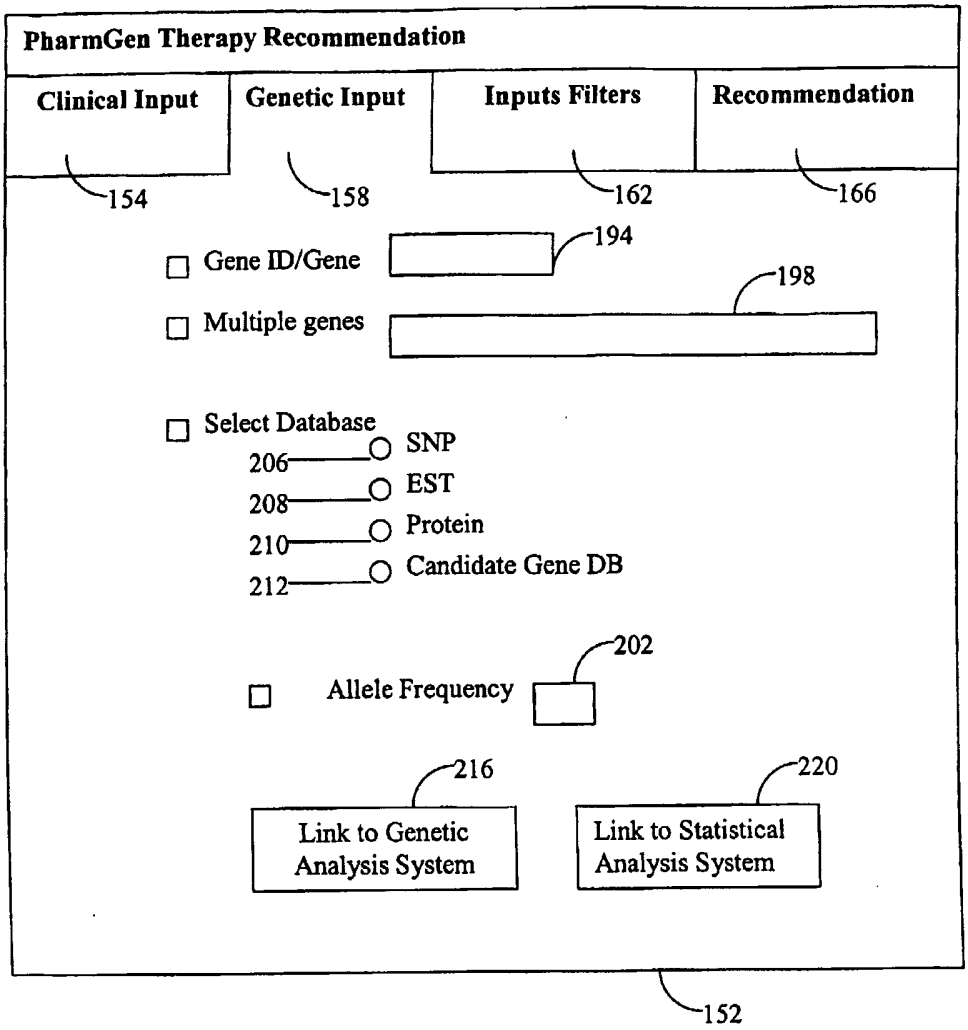
PharmGen Therapy Recommendation
Clinical Input
Genetic Input
Inputs Filters
Recommendation
154
158
162
166
Gene ID/Gene
194
Multiple genes
198
Select Database
206
SNP
208
EST
210
Protein
212
Candidate Gene DB
Allele Frequency
202
216
Link to Genetic Analysis System
220
Link to Statistical Analysis System
152

Fig. 5D

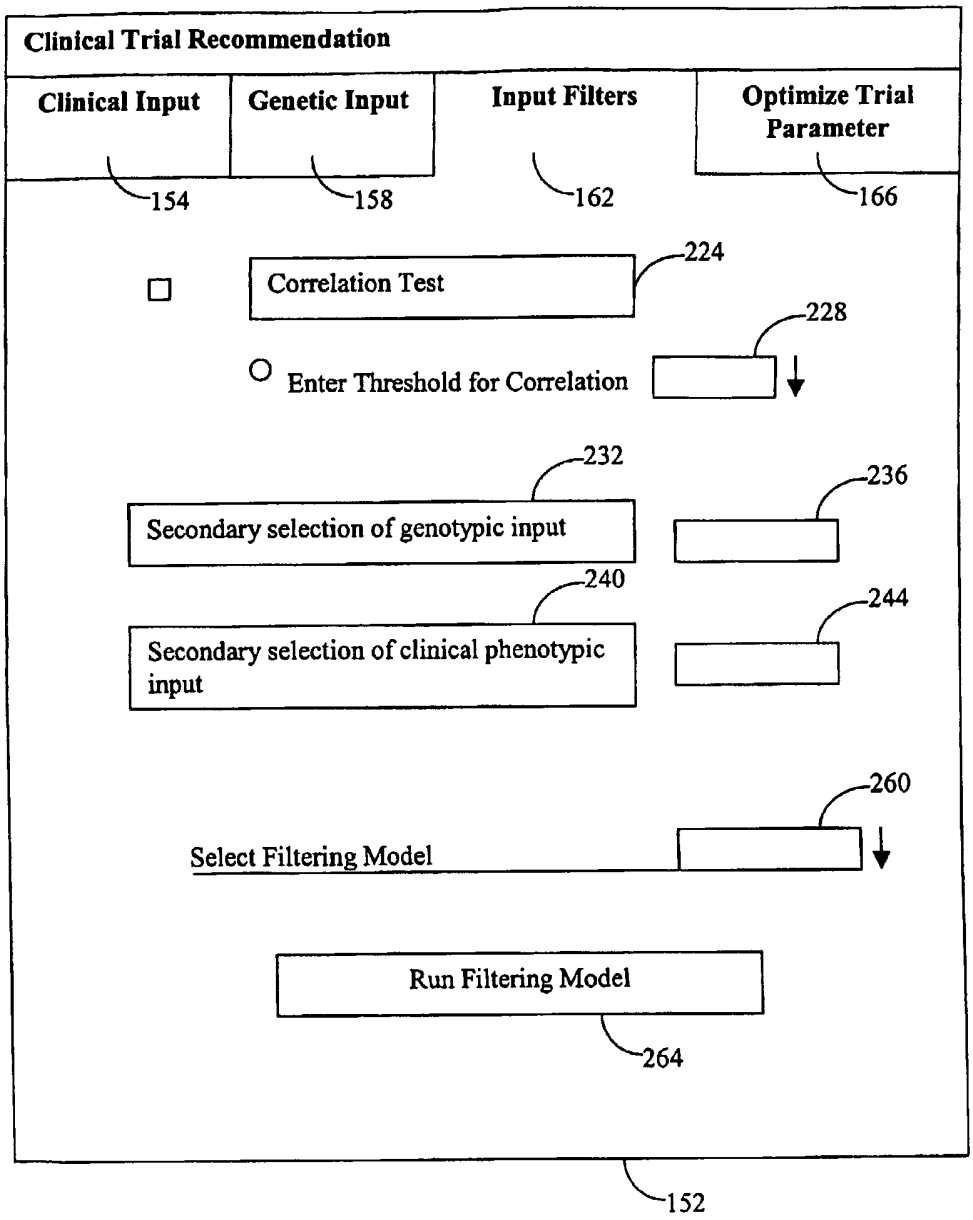
Clinical Trial Recommendation
Clinical Input
Genetic Input
Input Filters
Optimize Trial Parameter
154
158
162
166
Correlation Test
224
228
Enter Threshold for Correlation
232
236
Secondary selection of genotypic input
240
244
Secondary selection of clinical phenotypic input
260
Select Filtering Model
Run Filtering Model
264
152

Fig. 5E

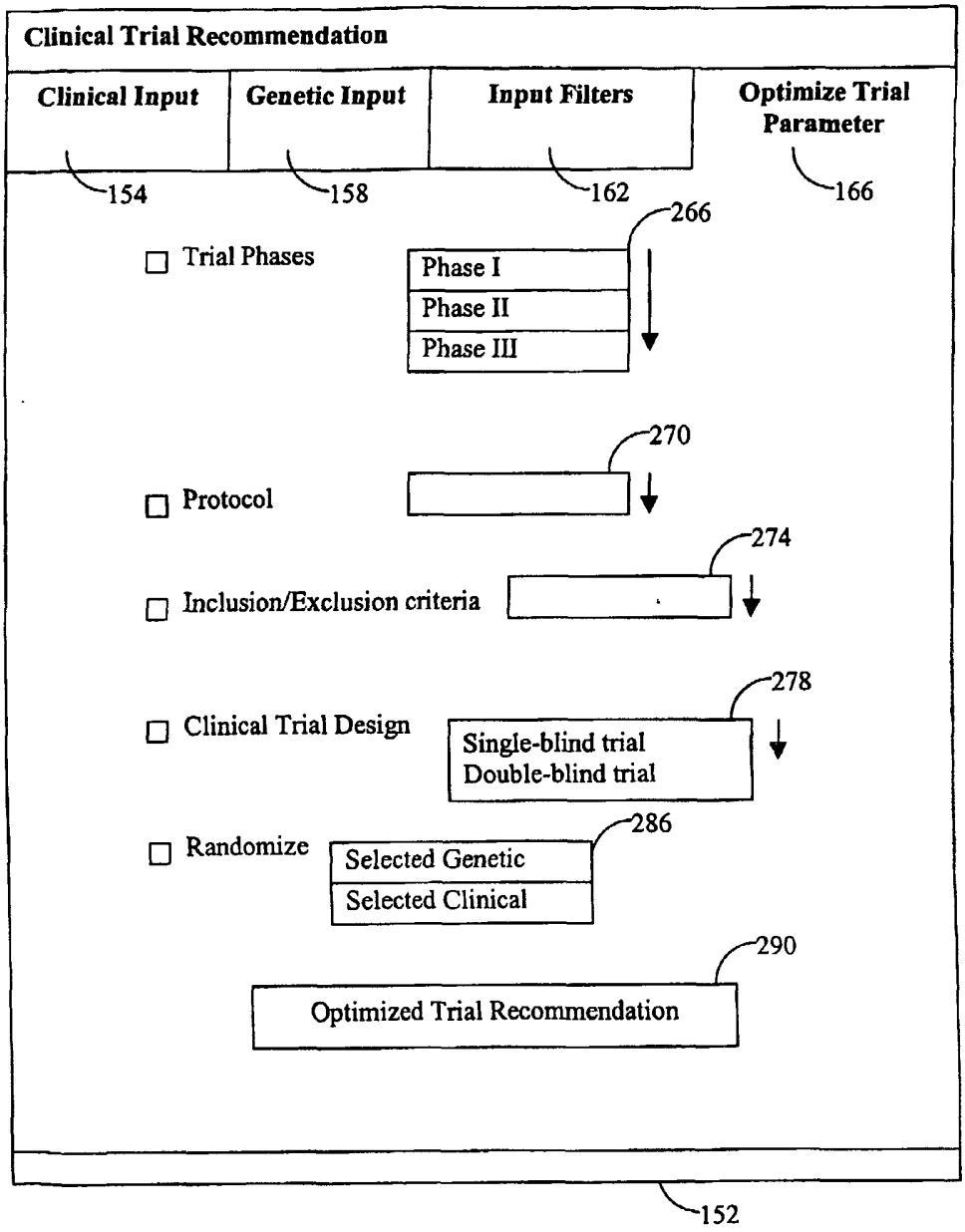
Clinical Trial Recommendation
Clinical Input
Genetic Input
Input Filters
Optimize Trial Parameter
154
158
162
166
266
Trial Phases
Phase I
Phase II
Phase III
270
Protocol
274
Inclusion/Exclusion criteria
278
Clinical Trial Design
Single-blind trial
Double-blind trial
Randomize
286
Selected Genetic
Selected Clinical
290
Optimized Trial Recommendation
152

Fig. 5F

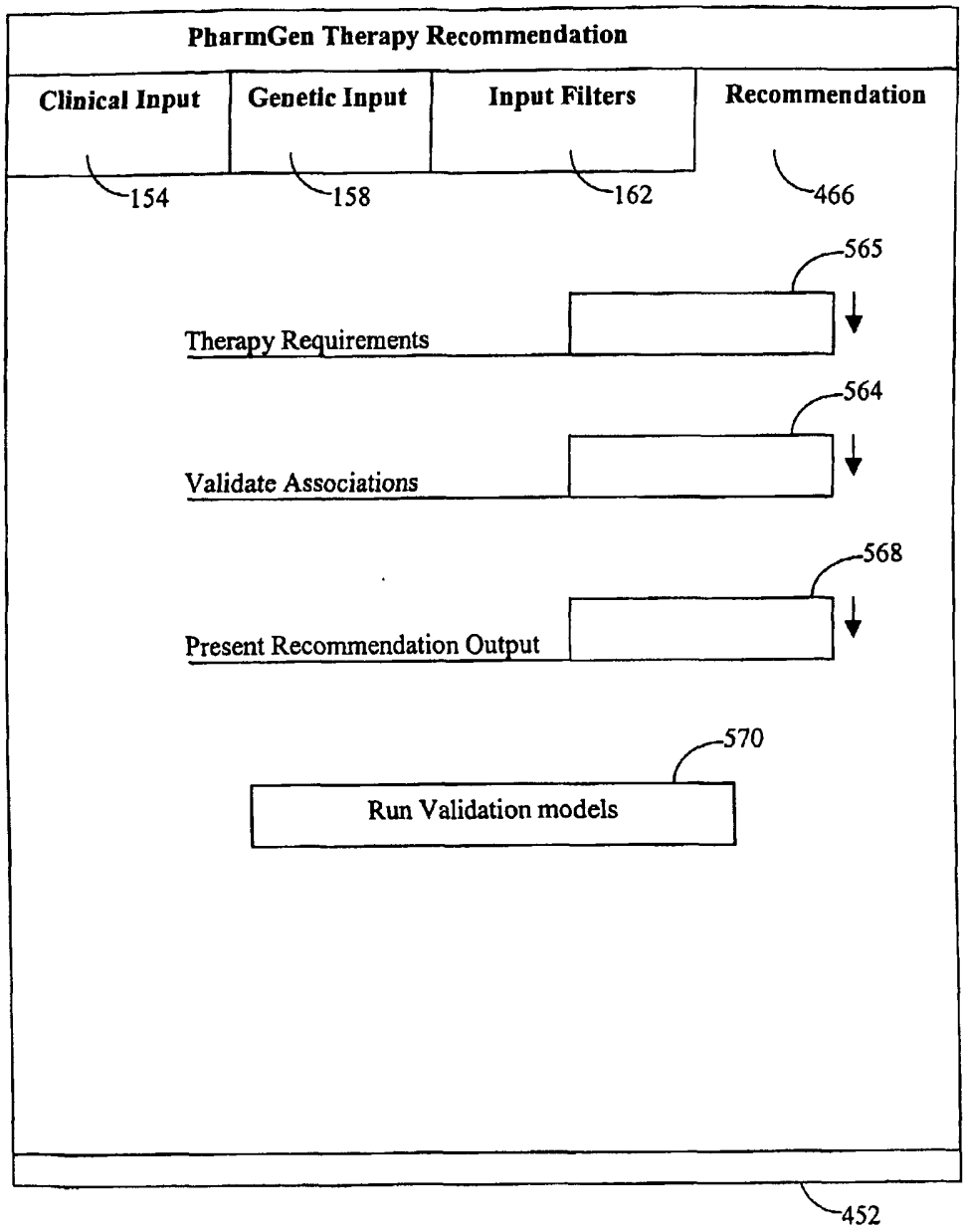
PharmGen Therapy Recommendation
Clinical Input
Genetic Input
Input Filters
Recommendation
154
158
162
466
565
Therapy Requirements
564
Validate Associations
568
Present Recommendation Output
570
Run Validation models
452

Fig. 5G

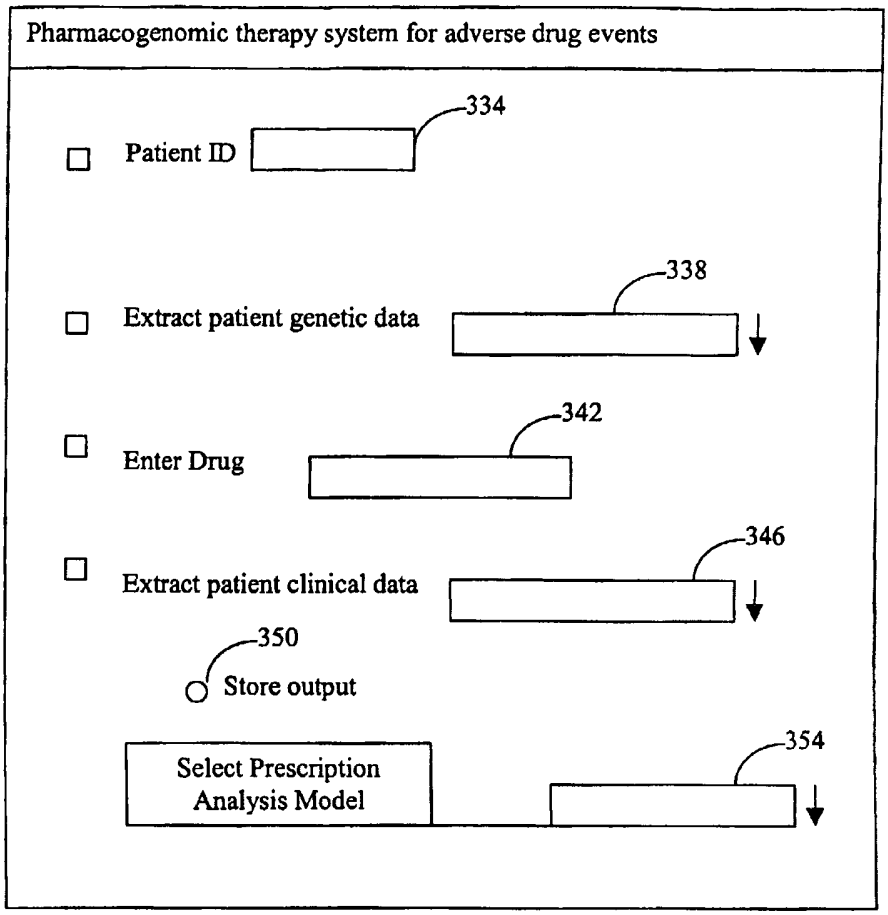
Pharmacogenomic therapy system for adverse drug events
Patient ID
334
Extract patient genetic data
338
Enter Drug
342
Extract patient clinical data
346
350
Store output
Select Prescription Analysis Model
354

Fig. 5H