

CLAIMS

What is claimed is:

1. A method implemented at one or more computing machines comprising processing circuitry and memory, the method comprising:

accessing, at the processing circuitry and from the memory, for a patient, a pharmacy transaction data set, each pharmacy transaction in the pharmacy transaction data set comprising at least a transaction date, a therapeutic type, and a therapeutic supply duration;

identifying, from the pharmacy transaction data set and based on the therapeutic type, a pharmacy transaction subset relevant to a biological condition;

computing, for at least one pharmacy transaction in the pharmacy transaction subset, an end date, wherein the transaction date corresponds to a date when the patient initiated a payment, wherein the end date is determined based on the transaction date and the therapeutic supply duration associated with the at least one pharmacy transaction;

accessing, at the processing circuitry and from the memory, for the patient, a medical procedure transaction data set, each medical procedure in the medical procedure data set comprising at least a medical procedure date range and a medical procedure type;

identifying, from the medical procedure transaction data set and based on the medical procedure type, a medical procedure transaction subset relevant to the biological condition;

adjusting, for at least one medical procedure in the medical procedure transaction subset, a medical procedure date range based on a time period during which the medical procedure type is effective or repeated;

mapping, by the processing circuitry, onto a timeline data structure stored in the memory, the pharmacy transaction subset and the medical procedure transaction subset, wherein the timeline data structure stores pharmacy transactions and medical procedure transactions arranged by date;

determining, by the processing circuitry, in the timeline data structure, one or more therapy gaps during which there are no pharmacy transactions and no medical procedure transactions, wherein each therapy gap comprises a contiguous number of days, the contiguous number of days being longer than a threshold number of days;

determining, by the processing circuitry, based on the one or more therapy gaps, one or more lines of therapy, each line of therapy comprising pharmacy transactions and medical procedure transactions occurring either between two therapy gaps, before an earliest temporal therapy gap, or after a latest temporal therapy gap, each line of therapy being associated with a line date range; and

transmitting, to a data repository for storage thereat, a data structure identifying the patient, the one or more lines of therapy, and the line date range for each of the one or more lines of therapy.

2. The method of claim 1, further comprising:

determining, within a single line of therapy, a first pharmacy transaction or medical procedure associated with a first biological condition stage and a second pharmacy transaction or medical procedure associated with a second biological condition stage, wherein a start date associated with the second pharmacy transaction or medical procedure is later than a start date associated with the first pharmacy transaction or medical procedure; and

dividing the single line of therapy into two lines of therapy using the start date associated with the second pharmacy transaction.

3. The method of any of claims 1-2, wherein:

the processing circuitry comprises a plurality of multithreaded graphics processing units (GPUs),

the patient is one of multiple patients,

the one or more lines of therapy for the patient are determined in parallel, using parallel threads of the plurality of multithreaded GPUs, with determining lines of therapy for other patients from among the multiple patients.

4. The method of claim 3, wherein determining the lines of therapy for the multiple patients comprises generating multiple intermediate tables, wherein each of the multiple intermediate tables is stored in the data repository for reviewing and adjusting the performance of the one or more computing machines.

5. The method of claim 4, wherein the multiple intermediate tables increase a computing speed of determining the lines of therapy for the multiple patients by storing intermediate computation results.

6. The method of any of claims 1-5, wherein, for at least one medical procedure transaction in the medical procedure transaction subset, the medical procedure date range comprises a single date, wherein the at least one medical procedure transaction is mapped onto the timeline data structure based on the single date.

7. The method of any of claims 1-6, wherein, for at least one medical procedure transaction in the medical procedure transaction subset, the medical procedure date range comprises a medical procedure start date and a medical procedure end date, wherein the at least one medical procedure transaction is mapped onto the timeline data structure based on the medical procedure start date and the medical procedure end date.

8. The method of any of claims 1-7, wherein pharmacy transactions from the pharmacy transaction subset are mapped onto the timeline data structure based on the transaction date.

9. The method of claim 8, wherein the at least one pharmacy transaction is mapped onto the timeline data structure based on the transaction date and the end date.

10. The method of any of claims 1-9, wherein the therapeutic type comprises a National Drug Code (NDC) classification, wherein the pharmacy transaction dataset comprises one or more tables, the method further comprising:

- determining a set of columns in the one or more tables related to drugs;
- parsing the set of columns to identify NDC classifications;
- determining a set of the NDC classifications that correspond to drugs;
- identifying a subset of the set of NDC classifications that correspond to drugs, the subset being associated with drugs related to the biological condition; and
- identifying rows in the one or more tables for placement into the pharmacy transaction

subset based on the group of NDC classifications related to the biological condition.

11. The method of any of claims 1-10, wherein the threshold number of days is determined based on a condition type of the biological condition.

12. A method comprising:

analyzing, by a computing system having one or more processors and memory, health insurance claims data stored by an integrated data repository to determine a number of patients that are at least one of diagnosed with a biological condition or have one or more genomic mutations detected, wherein the integrated data repository stores the health insurance claims data in conjunction with genomic data of the number of patients;

determining, by the computing system, a subset of the health insurance claims data that corresponds to the number of patients;

analyzing, by the computing system, the subset of the health insurance claims data to determine medical procedures records for first patients of the number of patients, the medical procedures records indicating one or more medical procedures administered to treat the biological condition in the first patients and indicating one or more first dates of the one or more medical procedures;

analyzing, by the computing system, the subset of health insurance claims data to determine pharmacy records for second patients of the number of patients, the pharmacy records indicating one or more pharmaceutical treatments provided to the second patients to treat the biological condition and indicating one or more second dates related to the one or more pharmaceutical treatments;

determining, by the computing system, a line of therapy corresponding to a patient of the number of patients based on at least one of the medical procedure records of the patient or the pharmacy records of the patient, the line of therapy indicating at least one of (i) a medical procedure administered to the patient to treat the biological condition and a first period of time that the medical procedure was administered or (ii) a pharmaceutical treatment provided to the individual to treat the biological condition and a second period of time that the pharmaceutical

treatment is to be provided to the patient;

generating, by the computing system, a lines of therapy data structure that includes the line of therapy of the patient and a number of additional lines of therapy of the first patients and the second patients;

analyzing, by the computing system, information stored by the lines of therapy data structure to determine one or more quantitative measures for the first patients and the second patients in relation to at least one of a medical procedure included in the one or more medical procedures or a pharmaceutical treatment of the one or more pharmaceutical treatments, the one or more quantitative measures corresponding to at least one of a first probability of progression of the biological condition or a second probability of resistance to at least one of the medical procedure or the pharmaceutical treatment; and

determining, by the computing system and based on the one or more quantitative measures, at least one of an amount of progression of the biological condition or a level of resistance to the pharmaceutical treatment for at least one of the first patients or the second patients.

13. The method of claim 12, comprising:

determining, by the computing system and based on the one or more quantitative measures, a level of effectiveness of at least one of the medical procedure or the pharmaceutical treatment for at least one of the first patients or the second patients.

14. The method of claim 12 or 13, wherein:

the pharmacy records indicate a first instance of the patient receiving a first supply of the pharmaceutical substance and a second instance of the patient receiving a second supply of the pharmaceutical substance;

the first supply of the pharmaceutical substance and the second supply of the pharmaceutical substance is for a number of days; and

the method comprises determining, by the computing system, that the first instance of the patient receiving the first supply takes place on a first date and the second instance of the patient receiving the second supply takes place on a second date.

15. The method of claim 14, comprising:

determining, by the computing system, that the second date is less than a threshold period of time after an intermediate date that comprises the number of days of the first supply added to the first date; and

determining, by the computing system, that the second instance of the patient receiving the second supply of the pharmaceutical substance is part of the line of therapy.

16. The method of claim 15, comprising:

determining, by the computing system, that the second date is at least a threshold period of time after an intermediate date that comprises the number of days of the first supply added to the first date; and

determining, by the computing system, that the second instance of the patient receiving the second supply of the pharmaceutical substance is part of an additional line of therapy that is different from the line of therapy.

17. The method of any one of claims 12-16, wherein:

the pharmacy records indicate a first instance of the patient receiving a first supply of the pharmaceutical substance and a second instance of the patient receiving a second supply of a second pharmaceutical substance;

the first supply of the pharmaceutical substance and the second supply of the pharmaceutical substance is for a number of days; and

the method comprises determining, by the computing system, that the first instance of the patient receiving the first supply takes place on a first date and the second instance of the patient receiving the second supply takes place on a second date.

18. The method of claim 17, comprising:

determining, by the computing system, that the second date is at least a threshold period of time after an intermediate date that comprises the number of days of the first supply added to the first date; and

determining, by the computing system, that the second instance of the patient receiving the

second supply of the second pharmaceutical substance is part of an additional line of therapy that is different from the line of therapy.

19. The method of claim 17, comprising:

determining, by the computing system, that the second date is at least a threshold period of time after an intermediate date that comprises the number of days of the first supply added to the first date;

determining, by the computing system, that the first pharmaceutical substance and the second pharmaceutical substance are part of a primary course of treatment for the biological condition; and

determining, by the computing system, that the second instance of the patient receiving the second supply of the second pharmaceutical substance is part of the line of therapy.

20. The method of any one of claims 12-19, wherein the one or more quantitative measures include at least one of time-to-next-treatment, time to death, or real world overall survival.

21. The method of claim 20, comprising:

determining, by the computing system, at least one of the amount of progression of the biological condition or the level of resistance to the pharmaceutical substance for the patient based on the time-to-next-treatment for the patient over a period of time with respect to the line of therapy and one or more additional lines of therapy.

22. The method of any one of claims 12-21, comprising:

generating, by the computing system, a procedures data table that stores the medical procedures records for first patients;

generating, by the computing system, a pharmacy data table storing the pharmacy records for second patients; and

analyzing, by the computing system, first information stored by the procedures data table and second information stored by the pharmacy data table according to a framework that indicates at least one of one or more threshold time periods, one or more criteria, or one or more data analysis

rules that correspond to the biological condition to generate a lines of therapy data structure that includes the line of therapy.

23. A system comprising:

one or more hardware processing units;

one or more computer-readable storage media storing computer-executable instructions that, when executed by the one or more hardware processing units, cause the system to perform the method of any one of the preceding claims.

24. One or more non-transitory computer-readable storage media storing computer-executable instructions that, when executed by one or more hardware processing units, cause a system to perform the method of any one of claims 1 to 22.