WHAT IS CLAIMED IS:

1. A computer-implemented method for determining a drug code, comprising:

acquiring, by one or more processors, a description text of a drug;

extracting, by the one or more processors using a natural language processing model, key information of the drug in the description text;

obtaining, by the one or more processors, at least one code related to the key information of the drug and an ingredient corresponding to each code by querying a pre-created code inverted index via a search engine; and

screening, by the one or more processors, the at least one code based on the key information of the drug and the ingredient corresponding to each code to obtain an anatomical therapeutic chemical classification system code of the drug.

2. The method according to claim 1, wherein the key information of the drug comprises a drug ingredient, and

screening the at least one code based on the key information of the drug and the ingredient corresponding to each code to obtain the anatomical therapeutic chemical classification system code of the drug comprises:

detecting, for each code in the at least one code, whether an ingredient corresponding to the code satisfies one of a plurality of rules having a priority order, the plurality of rules being determined based on the drug ingredient;

determining a preliminarily screened candidate code comprising the code, in response to determining that the ingredient corresponding to the code satisfies one of the plurality of rules and detections for all codes are completed; and

determining that the preliminarily screened candidate code is the anatomical therapeutic chemical classification system code of the drug, in response to detecting that the preliminarily screened candidate code refers to only one code.

3. The method according to claim 2, wherein the plurality of rules are arranged in a descending order of priorities as follows:

(1) the ingredient corresponding to the code comprises all drug ingredients of the drug when there are two or more kinds of drug ingredients;

(2) the ingredient corresponding to the code comprises at least one drug ingredient of the drug and contains a word “compound” when there are two or more kinds of drug ingredients;

(3) the ingredient corresponding to the code comprises at least one drug ingredient of the drug and does not contain the word “compound” when there are two or more kinds of drug ingredients; and

(4) the ingredient corresponding to the code comprises the drug ingredient when there is one kind of drug ingredient.

4. The method according to claim 1, wherein the key information of the drug comprises a drug ingredient, and

screening the at least one code based on the key information of the drug and the ingredient corresponding to each code to obtain the anatomical therapeutic chemical classification system code of the drug comprises:

detecting, for each code in the at least one code, whether an ingredient corresponding to the code matches the drug ingredient;

obtaining a preliminarily screened candidate code comprising the code, in response to determining that the ingredient corresponding to the code matches the drug ingredient and detections for all codes are completed; and

determining that the preliminarily screened candidate code is the anatomical therapeutic chemical classification system code of the drug, in response to detecting that the preliminarily screened candidate code refers to only one code.

5. The method according to any one of claims 2-4, wherein the key information of the drug further comprises a drug indication, and

screening the at least one code based on the key information of the drug and the ingredient corresponding to each code to obtain the anatomical therapeutic chemical classification system code of the drug further comprises:

determining a disease type corresponding to the drug based on the drug indication, in response to detecting that the preliminarily screened candidate code refers to a plurality of codes; and

screening a code corresponding to the disease type from the preliminarily

screened candidate code as the anatomical therapeutic chemical classification system code of the drug.

6. The method according to claim 5, wherein determining the disease type corresponding to the drug based on the drug indication comprises:

performing a disease classification on the indication using a pre-trained classification model to obtain the disease type outputted by the classification model.

7. The method according to any one of claims 1-6, further comprising:

obtaining ATC classification information comprising ATC codes and corresponding drug ingredients; and

performing inverse indexing on the ATC classification information using a search engine software to create the pre-created code inverted index, which represents an association from an English and/or Chinese drug name to a corresponding ATC code.

8. An apparatus for determining a drug code, comprising:

an acquiring unit, configured to acquire a description text of a drug;

an extracting unit, configured to extract key information of the drug in the description text;

an obtaining unit, configured to obtain at least one code related to the key information of the drug and an ingredient corresponding to each code by querying a pre-created code inverted index via a search engine; and

a screening unit, configured to screen the at least one code based on the key information of the drug and the ingredient corresponding to each code to obtain an anatomical therapeutic chemical classification system code of the drug.

9. The apparatus according to claim 8, wherein the key information of the drug comprises a drug ingredient, and the screening unit comprises:

a detecting module, configured to detect, for each code in the at least one code, whether an ingredient corresponding to the code satisfies one of a plurality of rules having a priority order, the plurality of rules being determined based on the drug ingredient;

a preliminarily screening module, configured to determine a preliminarily screened candidate code comprising the code, in response to determining that the ingredient corresponding to the code satisfies one of the plurality of rules and detections for all codes are completed; and

a determining module, configured to determine that the preliminarily screened candidate code is the anatomical therapeutic chemical classification system code of the drug, in response to detecting that the preliminarily screened candidate code refers to only one code.

10. The apparatus according to claim 9, wherein the plurality of rules are arranged in a descending order of priorities as follows:

(1) the ingredient corresponding to the code comprises all drug ingredients of the drug when there are two or more kinds of drug ingredients;

(2) the ingredient corresponding to the code comprises at least one drug ingredient of the drug and contains a word “compound” when there are two or more kinds of drug ingredients;

(3) the ingredient corresponding to the code comprises at least one drug ingredient of the drug and does not contain the word “compound” when there are two or more kinds of drug ingredients; and

(4) the ingredient corresponding to the code comprises the drug ingredient when there is one kind of drug ingredient.

11. The apparatus according to claim 8, wherein the key information of the drug comprises a drug ingredient,

and the screening unit further comprises:

a matching module configured to detect, for each code in the at least one code, whether an ingredient corresponding to the code matches the drug ingredient;

a responding module configured to obtain a preliminarily screened candidate code including the code, in response to determining that the ingredient corresponding to the code matches the drug ingredient and detections for all codes are completed; and

an encoding module configured to determine that the preliminarily screened candidate code is the anatomical therapeutic chemical classification system code of the drug, in response to detecting that the preliminarily screened candidate code refers to only one code.

12. The apparatus according to any one of claims 9-11 , wherein the key information of the drug further comprises a drug indication, and

the screening unit further comprises:

a classifying module, configured to determine a disease type corresponding to the drug based on the drug indication, in response to detecting that the preliminarily screened candidate code refers to a plurality of codes; and

a confirming module, configured to screen a code corresponding to the disease type from the preliminarily screened candidate code as the anatomical therapeutic chemical classification system code of the drug.

13. The apparatus according to claim 12, where the classifying module is further configured to perform a disease classification on the indication using a pre-trained classification model to obtain the disease type outputted by the classification model.

14. An electronic device, comprising:

one or more processors; and

a storage apparatus, storing one or more programs,

wherein the one or more programs, when executed by the one or more processors, cause the one or more processors to implement the method according to any one of claims 1-6.

15. A computer readable medium, storing a computer program, wherein the program, when executed by a processor, implements the method according to any one of claims 1-6.

16. A computer program product, comprising a computer program, wherein the computer program, when executed by a processor, implements the method according to any one of claims 1-6.