

DISEASE ABNORMAL DATA DETECTION METHOD, DEVICE, COMPUTING DEVICE, AND STORAGE MEDIUM

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority of Chinese Patent Application No. 201810321852.9, entitled “disease abnormal data detection method, device, computing device, and storage medium” filed on April 11, 2018 in the China National Intellectual Property Administration (CNIPA), the entire contents of which are incorporated by reference herein.

FIELD

The present disclosure relates to a technical field of data processing, specifically a disease abnormal data detection method, a device, a computing device, and a storage medium.

BACKGROUND

With the acceleration of the process of global economic integration, economic and exchange activities have increased, and the flow of people has become more frequent, providing a favorable environment for the spread and outbreak of diseases, and public health problems have become more and more

serious. At the same time, society and the natural environment are also undergoing changes. The increase in environmental pollution and natural disasters that affect public health has also increased the possibility of public health emergencies.

How to detect the abnormal data of diseases so as to be able to identify epidemics or public health emergencies early, and to take corresponding control measures as soon as possible to minimize the losses caused by public health emergencies, which has become an urgent problem to be solved.

Existing anomaly detection methods, such as z-score anomaly detection method and Grubbs anomaly detection method, require data to meet a normal distribution, but in fact, often fail in this requirement. For the traditional quartile method, all the data will be used, while the long past data has low reference value to the current data, and the results are more prone to deviation.

SUMMARY

In view of the above, it is necessary to propose a disease abnormal data detection method, a device, a computing device, and a non-transitory readable storage medium, which can realize efficient and accurate disease data abnormal detection.

A first aspect of the present disclosure provides a disease abnormal data detection method, the method includes:

forming time series data X by acquiring disease monitoring data from time point 0 to time point t , $X = [x_0, x_1, x_2, \dots, x_t]$;

setting a time window size w , for each time point i from time point w to time point t , calculating a mean μ_i and a standard deviation σ_i of the disease monitoring data in a time window corresponding to the time point i , and the size of the time window corresponding to each time point being w , $i = w, w + 1, \dots, t$;

obtaining an mz list $[mz_w, mz_{w+1}, mz_{w+2}, \dots, mz_t]$ by calculating a moving-z-score value of disease monitoring data x_i corresponding to the time point i , according to the mean μ_i and the standard deviation σ_i of the disease monitoring data corresponding to each time point i , where $mz_i = \frac{|x_i - \mu_i|}{\sigma_i}$;

determining an abnormal value threshold of the time series data X according to the mz list;

determining whether a moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than the abnormal value threshold, and when the moving-z-score value of disease monitoring data x_i corresponding to time point i is greater than the abnormal value threshold,

determining the disease monitoring data x_i corresponding to the time point i is an abnormal value.

A second aspect of the present disclosure provides a disease abnormal data detection device, the device includes:

an acquisition unit, configured to form time series data X by acquiring disease monitoring data from time point 0 to time point t , $X = [x_0, x_1, x_2, \dots, x_t]$;

a calculation unit, configured to set a time window size w , for each time point i from time point w to time point t , calculate a mean μ_i and a standard deviation σ_i of the disease monitoring data in a time window corresponding to the time point i , and the size of the time window corresponding to each time point being w , $i = w, w + 1, \dots, t$;

a calculation unit, also configured to obtain an mz list $[mz_w, mz_{w+1}, mz_{w+2}, \dots, mz_t]$ by calculating a moving-z-score value of disease monitoring data x_i corresponding to the time point i , according to the mean μ_i and the standard deviation σ_i of the disease monitoring data corresponding to each time point i , where $mz_i = \frac{|x_i - \mu_i|}{\sigma_i}$;

a determination unit, configured to determine an abnormal value threshold of the time series data X according to the mz list;

a comparing unit, configured to determine whether a moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than

the abnormal value threshold, and when the moving-z-score value of disease monitoring data x_i corresponding to time point i is greater than the abnormal value threshold, determine the disease monitoring data x_i corresponding to the time point i is an abnormal value.

A third aspect of the present disclosure provides a computing device, the computing device includes a processor and a storage device, and the storage device stores at least one computer-readable instruction, the processor executes the at least one computer-readable instruction to implement the disease abnormal data detection method.

A fourth aspect of the present disclosure provides a non-transitory readable storage medium having stored thereon at least one computer-readable instruction, and the at least one computer-readable instruction is executed by a processor to implement the disease abnormal data detection method.

The present disclosure forms time series data X by acquiring disease monitoring data from time point 0 to time point t , $X = [x_0, x_1, x_2, \dots, x_t]$; sets a time window size w , for each time point i from time point w to time point t , calculates a mean μ_i and a standard deviation σ_i of the disease monitoring data in a time window corresponding to the time point i , and the size of the time window corresponding to the time point is w , $i = w, w + 1, \dots, t$. An mz list $[mz_w, mz_{w+1}, mz_{w+2}, \dots, mz_t]$ is obtained by calculating a moving-z-score value of

disease monitoring data x_i corresponding to the time point i , according to the mean μ_i and the standard deviation σ_i of the disease monitoring data corresponding to each time point i , where $mz_i = \frac{|x_i - \mu_i|}{\sigma_i}$. An abnormal value threshold of the time series data X is determined according to the mz list. The method determines whether a moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than the abnormal value threshold, and when the moving-z-score value of disease monitoring data x_i corresponding to time point i is greater than the abnormal value threshold, determines that the disease monitoring data x_i corresponding to the time point i is an abnormal value.

The present disclosure introduces a time window and uses the historical neighboring data distribution with a most reference value as a reference, so the detection result is more reasonable. Due to a small amount of data in a time window, it is difficult for the data in the time window to meet a normal distribution, so a traditional z-score method cannot be used to detect abnormal values in the disease monitoring data. Embodiment one calculates the moving-z-score value of multiple disease monitoring data to ensure that an amount of the data is large enough to meet a condition of a normal distribution, and then determines the abnormal value threshold based on all moving-z-score values, thereby detecting the abnormal value in the disease monitoring data. Therefore,

Embodiment one has higher applicability for abnormal detection of disease monitoring data and realizes efficient and accurate abnormal detection of disease monitoring data.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows a flowchart a disease abnormal data detection method provided in Embodiment one of the present disclosure.

FIG. 2 shows a schematic structural diagram of a disease abnormal data detection device provided in Embodiment 2 of the present disclosure.

FIG. 3 shows a schematic structural diagram of a computing device provided in Embodiment 4 of the present disclosure.

DETAILED DESCRIPTION

For clarity of the illustration of objectives, features, and advantages of the present disclosure, the drawings combined with the detailed description illustrate the embodiments of the present disclosure hereinafter. It is noted that embodiments of the present disclosure and features of the embodiments can be combined, when there is no conflict.

Various details are described in the following descriptions for a better understanding of the present disclosure, however, the present disclosure may

also be implemented in other ways other than those described herein. The scope of the present disclosure is not to be limited by the specific embodiments disclosed below.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the present disclosure belongs. The terms used herein in the present disclosure are only for the purpose of describing specific embodiments, and are not intended to limit the present disclosure.

Optionally, the disease abnormal data detection method of the present disclosure is applied to one or more computing devices. The computing device is a device that can automatically perform numerical calculation and/or information processing according to pre-set or stored instructions, and its hardware includes, but is not limited to, a microprocessor and an Application Specific Integrated Circuit (ASIC), Field-Programmable Gate Array (FPGA), Digital Signal Processor (DSP), embedded devices, etc..

The computing device may be a device such as a desktop computer, a notebook, a palmtop computer, and a cloud server. The computer device can interact with users through a keyboard, a mouse, a remote control, a touch panel, or a voice control device.

Embodiment one

FIG. 1 is a flowchart of a disease abnormal data detection method in an embodiment of the present disclosure. The disease abnormal data detection method is applied to computing devices. The disease abnormal data detection method detects abnormal values in disease monitoring data, so as to be able to identify epidemics or public health emergencies early, and to take corresponding control measures as soon as possible to minimize the losses caused by public health emergencies.

As shown in Fig. 1, the disease abnormal data detection method specifically includes the following steps:

In block 101, forming time series data X by acquiring disease monitoring data from time point 0 to time point t , $X = [x_0, x_1, x_2, \dots, x_t]$.

The disease monitoring data may include monitoring data for diseases such as influenza, hand, foot and mouth disease, measles, mumps, etc..

A disease monitoring network composed of multiple monitoring points can be established in a preset area (for example, a province, a city, or a region), and the disease monitoring data can be acquired from the multiple monitoring points. Medical institutions, schools and childcare institutions, and pharmacies, etc. can be selected as the monitoring points to conduct disease monitoring and data collection for corresponding target populations. A place that meets preset conditions can be selected as the monitoring point. The preset conditions may include a number of

people, a level, and so on. For example, schools and kindergartens with a preset number of students can be selected as the monitoring points. For another example, a pharmacy with a level of operations (for example, counting daily turnover) reaches a preset level can be selected as the monitoring point. For another example, a hospital with a level of operations (for example, counting a number of medical treatments in a day) reaching a preset level can be selected as the monitoring point.

Disease monitoring data at different times form the time series data of disease monitoring. For example, disease monitoring data collected in days can form the time series data of disease monitoring. Alternatively, disease monitoring data collected in weeks can form the time series data of disease monitoring.

Medical institutions (mainly including hospitals) are the best places to capture early signs of disease outbreaks and are the first choice for disease monitoring. The disease monitoring data can be acquired according to a patient's medical situation.

Some people with diseases will go to pharmacies to purchase medicines to relieve early symptoms. Therefore, the disease monitoring data can be acquired according to the sales of medicines in the pharmacies.

Children and adolescents are high-risk groups in relation to diseases and important links in a process of disease transmission. The monitoring of these groups should also be strengthened. Schools and kindergartens are better places to

monitor the incidence of diseases in children and adolescents. The disease monitoring data can be acquired according to the absence-through-sickness status of children and adolescents in schools and kindergartens.

Therefore, in the present disclosure, three types of places, namely medical institutions, schools, childcare institutions, and pharmacies, are mainly selected to collect disease monitoring data. Of course, the above-mentioned selection of data sources cannot be limited to adding or replacing other people or places that focus on monitoring as data sources for monitoring in other embodiments. For example, a hotel can be included in a scope of disease monitoring to obtain disease monitoring data of hotel residents.

According to needs, disease monitoring data collected by any type of monitoring points (such as a medical institution) can be used to form the time series data of disease monitoring. For example, disease monitoring data collected by a hospital can be used to form the time series data of disease monitoring. Alternatively, disease monitoring data collected by multiple types of monitoring points can be combined to form the time series data of disease monitoring. For example, disease monitoring data collected by the hospital can be a mainstay, supplemented by disease monitoring data participated by the pharmacy to form the time series data of disease monitoring.

Disease monitoring data may include disease data such as a number of consultations, a consultation rate, a number of infections, and an incidence rate of the disease. For example, a number of daily consultations of diseases (such as influenza) can be obtained from medical institutions (such as hospitals), and the number of daily consultations for diseases (such as influenza) can be used as disease monitoring data. For another example, a number of daily infections of diseases (such as influenza) of students can be obtained from schools, and the number of daily infections of diseases (such as influenza) can be used as disease monitoring data.

It should be noted that the time point 0 represents a start time and the time point t represents an end time of the time series data, neither is time-limiting on the time series data. Any time point can be regarded as the time point 0.

In block 102, setting a time window size w , for each time point i from time point w to time point t , calculating a mean μ_i and a standard deviation σ_i of disease monitoring data in a time window corresponding to the time point i , a size of the time window corresponding to the time point i being w , $i = w, w + 1, \dots, t$.

The time window is an adjacent period before a given time point. For example, set the time window size w to 4, the given time point is $t=10$, and the time window is a period from $t=6$ to $t=9$; the given time point is $t=11$, and the time window is a period from $t=7$ to $t=11$; and so on.

The time window size w is an adjustable parameter, which can be adjusted according to actual conditions.

In an embodiment, the disease monitoring data is daily disease monitoring data (for example, a number of daily consultations of diseases), and the time window size w can be taken as 7 (ie, one week).

In another embodiment, the disease monitoring data is weekly disease monitoring data (for example, a number of weekly consultations of diseases), and the time window size w can be taken as 3 (ie, three weeks).

In block 103, obtaining an mz list $[mz_w, mz_{w+1}, mz_{w+2}, \dots, mz_t]$ by calculating a moving-z-score value mz_i of disease monitoring data x_i corresponding to the time point i , according to the mean μ_i and the standard deviation σ_i of the disease monitoring data corresponding to each time point i , where $mz_i = \frac{|x_i - \mu_i|}{\sigma_i}$.

The blocks 102-103 calculate the mean value and the standard deviation of the disease monitoring data in the time window corresponding to each time point from the time point w , and calculate the moving-z-score value of the disease monitoring data corresponding to each time point from the time point w . This is because a length of a time window obtained from $t=0$ to $t=w-1$ is less than w , and there is not enough w disease monitoring data to calculate the mean and the standard deviation of disease monitoring data.

In block 104, determining an abnormal value threshold of the time series data X according to the mz list.

A preset quantile or a preset percentile of the mz list may be calculated, and the preset quantile or the preset percentile can be used as the abnormal value thresholds. For example, a 0.997th quantile of the mz list is calculated, and the 0.997th quantile of the mz list is used as the abnormal value threshold. For another example, a 0.5th quantile (that is, a median) of the mz list is calculated, and the 0.5th quantile (that is, the median) of the mz list is used as the abnormal value threshold.

Alternatively, an average number of the mz list can be calculated, and the average number can be used as the abnormal value threshold.

In block 105, determining whether a moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than the abnormal value thresholds. When the moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than the abnormal value thresholds, determining that the disease monitoring data x_i corresponding to the time point i is an abnormal value.

For example, an abnormal value threshold is 12, and a moving-z-score value of a disease monitoring data x_9 corresponding to a time point 9 is 13, which is

greater than the abnormal value threshold, and the disease monitoring data x_9 is thus an abnormal value.

for example:

Disease monitoring data at 29 consecutive time points are:

[3,3,4,5,4,6,7,6,7,8,9,8,17,8,7,9,10,11,12,12,13,12,14,14,15,16,16,17,16].

Taking a time window size $w=4$, that is, calculating a moving-z-score value from a fifth number 4. For the fifth number, disease monitoring data in the time window is [3,3,4,5], an average value is 3.75 and a standard deviation is 0.829. A moving-z-score value is $(4-3.75)/0.829=0.302$ (rounded to three decimal places).

In the same way, calculating all moving-z-score values after the fifth number in turn to obtain an mz list:

[0.302, 2.828, 2.713, 0.447, 1.147, 3.0, 2.828, 0.447, 12.728, 0.662, 0.927, 0.246, 0.063, 2.236, 1.859, 1.342, 2.110, 0.0, 4.041, 1.508, 2.110, 2.065, 1.508, 2.111, 0.0].

Then taking a 0.995 quantile of the mz list as the abnormal value threshold to obtain 11.685.

Only 12.728 in the mz list is greater than the abnormal value threshold, 12.728 corresponds to a disease monitoring data 17, so 17 is the abnormal value.

The disease abnormal data detection method in the Embodiment one forms time series data X by acquiring disease monitoring data from time point 0 to time

point t , $X = [x_0, x_1, x_2, \dots, x_t]$; sets a time window size w , for each time point i from time point w to time point t , calculates a mean μ_i and a standard deviation σ_i of the disease monitoring data in a time window corresponding to the time point i , and the size of the time window corresponding to the time point is w , $i = w, w + 1, \dots, t$. An mz list $[mz_w, mz_{w+1}, mz_{w+2}, \dots, mz_t]$ is obtained by calculating a moving-z-score value of disease monitoring data x_i corresponding to the time point i , according to the mean μ_i and the standard deviation σ_i of the disease monitoring data corresponding to each time point i , where $mz_i = \frac{|x_i - \mu_i|}{\sigma_i}$. An abnormal value threshold of the time series data X is determined according to the mz list. The method determines whether a moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than the abnormal value threshold, and when the moving-z-score value of disease monitoring data x_i corresponding to time point i is greater than the abnormal value threshold, determines that the disease monitoring data x_i corresponding to the time point i is an abnormal value.

Embodiment one introduces a time window and uses the historical neighboring data distribution with a most reference value as a reference, so the detection result is more reasonable. Due to a small amount of data in a time window, it is difficult for the data in the time window to meet a normal distribution, so a traditional z-score method cannot be used to detect abnormal values in the

disease monitoring data. Embodiment one calculates the moving-z-score value of multiple disease monitoring data to ensure that an amount of the data is large enough to meet a condition of a normal distribution, and then determines the abnormal value threshold based on all moving-z-score values, thereby detecting the abnormal value in the disease monitoring data. Therefore, Embodiment one has higher applicability for abnormal detection of disease monitoring data and realizes efficient and accurate abnormal detection of disease monitoring data.

In another embodiment, other disease abnormal data detection methods (different from the disease abnormal data detection method in Embodiment one) can be combined to perform disease abnormal data detection. Specifically, it can further include the following steps: obtaining abnormal values obtained by detecting the time series data by other disease abnormal data detection methods; comparing the abnormal values obtained by the disease abnormal data detection method in Embodiment one with the abnormal values obtained by the other disease abnormal data detection method; obtaining a final abnormal value according to a comparison result of the abnormal value obtained by the disease abnormal data detection method in Embodiment one and the abnormal values obtained by the other disease abnormal data detection method. The other disease abnormal data detection methods can a single method or multiple methods. Every other disease abnormal data detection method can obtain a set of corresponding abnormal values.

For example, using the disease abnormal data detection method in Embodiment one to perform disease abnormal data detection on the time series data of disease monitoring to obtain a first abnormal value, using a second disease abnormal data detection method (such as Grubbs detection method) different from the disease abnormal data detection method in Embodiment one to perform disease abnormal data detection on the time series data of disease monitoring to obtain a second abnormal value, and comparing the first abnormal value with the second abnormal value to obtain a final abnormal value. When the first abnormal value is consistent with the second abnormal value, the first abnormal value/the second abnormal value is taken as the final abnormal value; otherwise, when the first abnormal value and the second abnormal value is inconsistent, the first abnormal value and the second abnormal value are discarded.

For another example, using the disease abnormal data detection method in Embodiment one to perform disease abnormal data detection on the time series data of disease monitoring to obtain a first abnormal value, using a second disease abnormal data detection method (for example Grubbs detection method) different from the disease abnormal data detection method in Embodiment one to perform disease abnormal data detection on the time series data of disease monitoring to obtain a second abnormal value, using a third disease abnormal data detection method (such as Bayesian detection method) different from the disease abnormal

data detection method in Embodiment one to perform disease abnormal data detection on the time series data of disease monitoring to obtain a third abnormal value, and comparing the first abnormal value, the second abnormal value, and the third abnormal value to obtain a final abnormal value. When there is consistency in at least two of the first abnormal value, the second abnormal value, and the third abnormal value, the consistent abnormal value is taken as the final abnormal value.

Based on the disease abnormal data detection method in Embodiment one, combined with other disease abnormal data detection methods for disease abnormal data detection, more accurate disease abnormal data detection results can be obtained.

In another embodiment, the disease abnormal data detection method in Embodiment one can be used to detect the disease abnormal data of regions and regional subordinate hospitals, and obtain detection results for the regions (that is, regional abnormal values) and detection results for the regional subordinate hospitals (that is, regional subordinate hospitals abnormal values). The regional abnormal values and the regional subordinate hospitals abnormal values is compared, and a final abnormal value is obtained according to the comparison result of the regional abnormal values and the regional subordinate hospitals abnormal values. If the detection results for the regions are the same as the detection results for the regional subordinate hospitals, the detection results for the

regions/the detection results for the regional subordinate hospitals will be the final detection results. Otherwise, if the detection results for the regions are not the same as the detection results for the regional subordinate hospitals, the detection results for the regions and the detection results for the regional subordinate hospitals are discarded.

The disease abnormal data detection for regions is to obtain the regional abnormal values based on time series data of disease monitoring in the regions (for example, the disease monitoring data collected by a disease control department of the region from various regional subordinate hospitals). The disease abnormal data detection for regional subordinate hospitals is to obtain the regional subordinate hospital abnormal values based on time series data of disease monitoring in the regional subordinate hospitals.

Specifically, using the disease abnormal data detection method in Embodiment one to obtain abnormal values from the disease monitoring data (such as a number of cases) in a certain period of time in the region, and to obtain abnormal values from the disease monitoring data in the regional subordinate hospitals in the same period of time. An intersection of the two dimensions is taken as the final detection result.

For example, the number of cases in the region in 2014-3-3, 2014-3-4, 2014-3-5, 2014-3-6 are 160, 250, 170, 180, respectively, and the subordinate hospitals in

the region (for example, the region has a most cumulative number of medical visits for this disease). The number of cases during this period was 130, 180, 125, and 140, respectively. From a distribution of the number of cases in the region, a number of abnormal cases appeared on the day of 2014-3-4, and a number of abnormal cases also appeared in the subordinate hospitals on this day. So, it can be comprehensively determined that there was an abnormal phenomenon of disease visits on this day.

For another example, the number of cases in the region in 2014-3-3, 2014-3-4, 2014-3-5, 2014-3-6 are 160, 210, 170, 180, respectively, and the subordinate hospitals in the region (for example, the region has a most cumulative number of medical visits for this disease). The number of cases during this period was 130, 140, 125, 140, respectively. From a distribution of the number of cases in the region, a number of abnormal cases appeared on the day of 2014-3-4, but a number of abnormal cases does not appear in the subordinate hospitals on this day. This is because there are many hospitals in the region, and the number of patients in the region on this day is a sum of the number of cases from all the hospitals under it. Some hospitals had a small increase on this day, which is not abnormal in the hospital dimension. But a sum of all the small increases is abnormal in the regional dimension.

The disease abnormal data detection method in Embodiment one is used to detect disease abnormal data of the regions and the regional subordinate hospitals respectively, and the final detection results are obtained according to the detection results for the region and the detection results for the regional subordinate hospitals, more accurate disease abnormal data detection results can be obtained.

Embodiment two

FIG. 2 shows a schematic structural diagram of a disease abnormal data detection device provided in Embodiment 2 of the present disclosure. As shown in Figure 2, the disease abnormal data detection device 10 can include: an acquisition unit 201, a calculation unit 202, a determination unit 203, and a comparing unit 204.

The acquisition unit 201, is configured to form time series data X by acquiring disease monitoring data from time point 0 to time point t , $X = [x_0, x_1, x_2, \dots, x_t]$.

The disease monitoring data may include monitoring data for diseases such as influenza, hand, foot and mouth disease, measles, mumps, etc..

A disease monitoring network composed of multiple monitoring points can be established in a preset area (for example, a province, a city, or a region), and the

disease monitoring data can be acquired from the multiple monitoring points. Medical institutions, schools and childcare institutions, and pharmacies, etc. can be selected as the monitoring points to conduct disease monitoring and data collection for corresponding target populations. A place that meets preset conditions can be selected as the monitoring point. The preset conditions may include a number of people, a level, and so on. For example, schools and kindergartens with a preset number of students can be selected as the monitoring points. For another example, a pharmacy with a level of operations (for example, counting daily turnover) reaches a preset level can be selected as the monitoring point. For another example, a hospital with a level of operations (for example, counting a number of medical treatments in a day) reaching a preset level can be selected as the monitoring point.

Disease monitoring data at different times form the time series data of disease monitoring. For example, disease monitoring data collected in days can form the time series data of disease monitoring. Alternatively, disease monitoring data collected in weeks can form the time series data of disease monitoring.

Medical institutions (mainly including hospitals) are the best places to capture early signs of disease outbreaks and are the first choice for disease monitoring. The disease monitoring data can be acquired according to a patient's medical situation.

Some people with diseases will go to pharmacies to purchase medicines to relieve early symptoms. Therefore, the disease monitoring data can be acquired according to the sales of medicines in the pharmacies.

Children and adolescents are high-risk groups in relation to diseases and important links in a process of disease transmission. The monitoring of these groups should also be strengthened. Schools and kindergartens are better places to monitor the incidence of diseases in children and adolescents. The disease monitoring data can be acquired according to the absence-through-sickness status of children and adolescents in schools and kindergartens.

Therefore, in the present disclosure, three types of places, namely medical institutions, schools, childcare institutions, and pharmacies, are mainly selected to collect disease monitoring data. Of course, the above-mentioned selection of data sources cannot be limited to adding or replacing other people or places that focus on monitoring as data sources for monitoring in other embodiments. For example, a hotel can be included in a scope of disease monitoring to obtain disease monitoring data of hotel residents.

According to needs, disease monitoring data collected by any type of monitoring points (such as a medical institution) can be used to form the time series data of disease monitoring. For example, disease monitoring data collected by a hospital can be used to form the time series data of disease monitoring.

Alternatively, disease monitoring data collected by multiple types of monitoring points can be combined to form the time series data of disease monitoring. For example, disease monitoring data collected by the hospital can be a mainstay, supplemented by disease monitoring data participated by the pharmacy to form the time series data of disease monitoring.

Disease monitoring data may include disease data such as a number of consultations, a consultation rate, a number of infections, and an incidence rate of the disease. For example, a number of daily consultations of diseases (such as influenza) can be obtained from medical institutions (such as hospitals), and the number of daily consultations for diseases (such as influenza) can be used as disease monitoring data. For another example, a number of daily infections of diseases (such as influenza) of students can be obtained from schools, and the number of daily infections of diseases (such as influenza) can be used as disease monitoring data.

It should be noted that the time point 0 represents a start time and the time point t represents an end time of the time series data, neither is time-limiting on the time series data. Any time point can be regarded as the time point 0.

The calculation unit 202, is configured to set a time window size w , for each time point i from time point w to time point t , calculate a mean μ_i and a standard deviation σ_i of disease monitoring data in a time window corresponding to the time

point i , a size of the time window corresponding to the time point i being w ,
 $i = w, w + 1, \dots, t$;

The time window is an adjacent period before a given time point. For example, set the time window size w to 4, the given time point is $t=10$, and the time window is a period from $t=6$ to $t=9$; the given time point is $t=11$, and the time window is a period from $t=7$ to $t=11$; and so on.

The time window size w is an adjustable parameter, which can be adjusted according to actual conditions.

In an embodiment, the disease monitoring data is daily disease monitoring data (for example, a number of daily consultations of diseases), and the time window size w can be taken as 7 (ie, one week).

In another embodiment, the disease monitoring data is weekly disease monitoring data (for example, a number of weekly consultations of diseases), and the time window size w can be taken as 3 (ie, three weeks).

The calculation unit 202, is also configured to obtain an mz list $[mz_w, mz_{w+1}, mz_{w+2}, \dots, mz_t]$ by calculating a moving-z-score value mz_i of disease monitoring data x_i corresponding to the time point i , according to the mean μ_i and the standard deviation σ_i of the disease monitoring data corresponding to each time point i , where $mz_i = \frac{|x_i - \mu_i|}{\sigma_i}$,

The calculation unit 202 calculates the mean value and the standard deviation of the disease monitoring data in the time window corresponding to each time point from the time point w , and calculates the moving-z-score value of the disease monitoring data corresponding to each time point from the time point w . This is because a length of a time window obtained from $t=0$ to $t=w-1$ is less than w , and there is not enough w disease monitoring data to calculate the mean and the standard deviation of disease monitoring data.

The determination unit 203, is configured to determine an abnormal value threshold of the time series data X according to the mz list.

A preset quantile or a preset percentile of the mz list may be calculated, and the preset quantile or the preset percentile can be used as the abnormal value thresholds. For example, a 0.997th quantile of the mz list is calculated, and the 0.997th quantile of the mz list is used as the abnormal value threshold. For another example, a 0.5th quantile (that is, a median) of the mz list is calculated, and the 0.5th quantile (that is, the median) of the mz list is used as the abnormal value threshold.

Alternatively, an average number of the mz list can be calculated, and the average number can be used as the abnormal value threshold.

The comparing unit 204, is configured to determine whether a moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than

the abnormal value thresholds. When the moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than the abnormal value thresholds, determining that the disease monitoring data x_i corresponding to the time point i is an abnormal value.

For example, an abnormal value threshold is 12, and a moving-z-score value of a disease monitoring data x_9 corresponding to a time point 9 is 13, which is greater than the abnormal value threshold, and the disease monitoring data x_9 is thus an abnormal value.

for example:

Disease monitoring data at 29 consecutive time points are:

[3,3,4,5,4,6,7,6,7,8,9,8,17,8,7,9,10,11,12,12,13,12,14,14,15,16,16,17,16].

Taking a time window size $w=4$, that is, calculating a moving-z-score value from a fifth number 4. For the fifth number, disease monitoring data in the time window is [3,3,4,5], an average value is 3.75 and a standard deviation is 0.829. A moving-z-score value is $(4-3.75)/0.829=0.302$ (rounded to three decimal places).

In the same way, calculating all moving-z-score values after the fifth number in turn to obtain an mz list:

[0.302, 2.828, 2.713, 0.447, 1.147, 3.0, 2.828, 0.447, 12.728, 0.662, 0.927, 0.246, 0.063, 2.236, 1.859, 1.342, 2.110, 0.0, 4.041, 1.508, 2.110, 2.065, 1.508, 2.111, 0.0].

Then taking a 0.995 quantile of the mz list as the abnormal value threshold to obtain 11.685.

Only 12.728 in the mz list is greater than the abnormal value threshold, 12.728 corresponds to a disease monitoring data 17, so 17 is the abnormal value.

The disease abnormal data detection method in the Embodiment two forms time series data X by acquiring disease monitoring data from time point 0 to time point t , $X = [x_0, x_1, x_2, \dots, x_t]$; sets a time window size w , for each time point i from time point w to time point t , calculates a mean μ_i and a standard deviation σ_i of the disease monitoring data in a time window corresponding to the time point i , and the size of the time window corresponding to the time point is w , $i = w, w + 1, \dots, t$. A mz list $[mz_w, mz_{w+1}, mz_{w+2}, \dots, mz_t]$ is obtained by calculating a moving-z-score value of disease monitoring data x_i corresponding to the time point i , according to the mean μ_i and the standard deviation σ_i of the disease monitoring data corresponding to each time point i , where $mz_i = \frac{|x_i - \mu_i|}{\sigma_i}$. An abnormal value threshold of the time series data X is determined according to the mz list. The method determines whether a moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than the abnormal value threshold, and when the moving-z-score value of disease monitoring data x_i corresponding to time

point i is greater than the abnormal value threshold, determines that the disease monitoring data x_i corresponding to the time point i is an abnormal value.

Embodiment two introduces a time window and uses the historical neighboring data distribution with a most reference value as a reference, so the detection result is more reasonable. Due to a small amount of data in a time window, it is difficult for the data in the time window to meet a normal distribution, so a traditional z-score method cannot be used to detect abnormal values in the disease monitoring data. Embodiment two calculates the moving-z-score value of multiple disease monitoring data to ensure that an amount of the data is large enough to meet a condition of a normal distribution, and then determines the abnormal value threshold based on all moving-z-score values, thereby detecting the abnormal value in the disease monitoring data. Therefore, Embodiment two has higher applicability for abnormal detection of disease monitoring data and realizes efficient and accurate abnormal detection of disease monitoring data.

In another embodiment, other disease abnormal data detection methods (different from the disease abnormal data detection method in Embodiment two) can be combined to perform disease abnormal data detection. Specifically, it can further include the following steps: obtaining abnormal values obtained by detecting the time series data by other disease abnormal data detection methods; comparing the abnormal values obtained by the disease abnormal data detection

method in Embodiment two with the abnormal values obtained by the other disease abnormal data detection method; obtaining a final abnormal value according to a comparison result of the abnormal value obtained by the disease abnormal data detection method in Embodiment two and the abnormal values obtained by the other disease abnormal data detection method. The other disease abnormal data detection methods can a single method or multiple methods. Every other disease abnormal data detection method can obtain a set of corresponding abnormal values.

For example, using the disease abnormal data detection method in Embodiment two to perform disease abnormal data detection on the time series data of disease monitoring to obtain a first abnormal value, using a second disease abnormal data detection method (such as Grubbs detection method) different from the disease abnormal data detection method in Embodiment two to perform disease abnormal data detection on the time series data of disease monitoring to obtain a second abnormal value, and comparing the first abnormal value with the second abnormal value to obtain a final abnormal value. When the first abnormal value is consistent with the second abnormal value, the first abnormal value/the second abnormal value is taken as the final abnormal value; otherwise, when the first abnormal value and the second abnormal value is inconsistent, the first abnormal value and the second abnormal value are discarded.

For another example, using the disease abnormal data detection method in Embodiment two to perform disease abnormal data detection on the time series data of disease monitoring to obtain a first abnormal value, using a second disease abnormal data detection method (for example Grubbs detection method) different from the disease abnormal data detection method in Embodiment two to perform disease abnormal data detection on the time series data of disease monitoring to obtain a second abnormal value, using a third disease abnormal data detection method (such as Bayesian detection method) different from the disease abnormal data detection method in Embodiment two to perform disease abnormal data detection on the time series data of disease monitoring to obtain a third abnormal value, and comparing the first abnormal value, the second abnormal value, and the third abnormal value to obtain a final abnormal value. When there is consistency in at least two of the first abnormal value, the second abnormal value, and the third abnormal value, the consistent abnormal value is taken as the final abnormal value.

Based on the disease abnormal data detection method in Embodiment two, combined with other disease abnormal data detection methods for disease abnormal data detection, more accurate disease abnormal data detection results can be obtained.

In another embodiment, the disease abnormal data detection method in Embodiment two can be used to detect the disease abnormal data of regions and

regional subordinate hospitals, and obtain detection results for the regions (that is, regional abnormal values) and detection results for the regional subordinate hospitals (that is, regional subordinate hospitals abnormal values). The regional abnormal values and the regional subordinate hospitals abnormal values is compared, and a final abnormal value is obtained according to the comparison result of the regional abnormal values and the regional subordinate hospitals abnormal values. If the detection results for the regions are the same as the detection results for the regional subordinate hospitals, the detection results for the regions/the detection results for the regional subordinate hospitals will be the final detection results. Otherwise, if the detection results for the regions are not the same as the detection results for the regional subordinate hospitals, the detection results for the regions and the detection results for the regional subordinate hospitals are discarded.

The disease abnormal data detection for regions is to obtain the regional abnormal values based on time series data of disease monitoring in the regions (for example, the disease monitoring data collected by a disease control department of the region from various regional subordinate hospitals). The disease abnormal data detection for regional subordinate hospitals is to obtain the regional subordinate hospital abnormal values based on time series data of disease monitoring in the regional subordinate hospitals.

Specifically, using the disease abnormal data detection method in Embodiment two to obtain abnormal values from the disease monitoring data (such as a number of cases) in a certain period of time in the region, and to obtain abnormal values from the disease monitoring data in the regional subordinate hospitals in the same period of time. An intersection of the two dimensions is taken as the final detection result.

For example, the number of cases in the region in 2014-3-3, 2014-3-4, 2014-3-5, 2014-3-6 are 160, 250, 170, 180, respectively, and the subordinate hospitals in the region (for example, the region has a most cumulative number of medical visits for this disease). The number of cases during this period was 130, 180, 125, and 140, respectively. From a distribution of the number of cases in the region, a number of abnormal cases appeared on the day of 2014-3-4, and a number of abnormal cases also appeared in the subordinate hospitals on this day. So, it can be comprehensively determined that there was an abnormal phenomenon of disease visits on this day.

For another example, the number of cases in the region in 2014-3-3, 2014-3-4, 2014-3-5, 2014-3-6 are 160, 210, 170, 180, respectively, and the subordinate hospitals in the region (for example, the region has a most cumulative number of medical visits for this disease). The number of cases during this period was 130, 140, 125, 140, respectively. From a distribution of the number of cases in the

region, a number of abnormal cases appeared on the day of 2014-3-4, but a number of abnormal cases does not appear in the subordinate hospitals on this day. This is because there are many hospitals in the region, and the number of patients in the region on this day is a sum of the number of cases from all the hospitals under it. Some hospitals had a small increase on this day, which is not abnormal in the hospital dimension. But a sum of all the small increases is abnormal in the regional dimension.

The disease abnormal data detection method in Embodiment two is used to detect disease abnormal data of the regions and the regional subordinate hospitals respectively, and the final detection results are obtained according to the detection results for the region and the detection results for the regional subordinate hospitals, more accurate disease abnormal data detection results can be obtained.

Embodiment three

The embodiment provides a non-transitory readable storage medium having computer-readable instructions stored on thereof. When the computer-readable instructions are executed by a processor to implement the steps in the above-mentioned disease abnormal data detection method, such as steps 101-105 shown in Fig. 1:

In block 101, forming time series data X by acquiring disease monitoring data from time point 0 to time point t , $X = [x_0, x_1, x_2, \dots, x_t]$;

In block 102, setting a time window size w , for each time point i from time point w to time point t , calculating a mean μ_i and a standard deviation σ_i of disease monitoring data in a time window corresponding to the time point i , a size of the time window corresponding to the time point i being w , $i = w, w + 1, \dots, t$;

In block 103, obtaining an mz list $[mz_w, mz_{w+1}, mz_{w+2}, \dots, mz_t]$ by calculating a moving-z-score value mz_i of disease monitoring data x_i corresponding to the time point i , according to the mean μ_i and the standard deviation σ_i of the disease monitoring data corresponding to each time point i , where $mz_i = \frac{|x_i - \mu_i|}{\sigma_i}$;

In block 104, determining an abnormal value threshold of the time series data X according to the mz list;

In block 105, determining whether a moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than the abnormal value thresholds. When the moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than the abnormal value thresholds, determining that the disease monitoring data x_i corresponding to the time point i is an abnormal value.

Or, the computer-readable instruction is executed by the processor to realize the functions of each module/unit in the above-mentioned device embodiments, such as the units 201-204 in Fig. 2:

The acquisition unit 201, is configured to form time series data X by acquiring disease monitoring data from time point 0 to time point t ,

$$X = [x_0, x_1, x_2, \dots, x_t];$$

The calculation unit 202, is configured to set a time window size w , for each time point i from time point w to time point t , calculate a mean μ_i and a standard deviation σ_i of disease monitoring data in a time window corresponding to the time point i , a size of the time window corresponding to the time point i being w ,

$$i = w, w + 1, \dots, t; i = w, w + 1, \dots, t;$$

The calculation unit 202, is configured to obtain an mz list $[mz_w, mz_{w+1}, mz_{w+2}, \dots, mz_t]$ by calculating a moving-z-score value mz_i of disease monitoring data x_i corresponding to the time point i , according to the mean μ_i and the standard deviation σ_i of the disease monitoring data corresponding to each time point i , where $mz_i = \frac{|x_i - \mu_i|}{\sigma_i}$;

The determination unit 203 is configured to determine an abnormal value threshold of the time series data X according to the mz list;

The comparing unit 204 is configured to determine whether a moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than the abnormal value thresholds. When the moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than the abnormal value thresholds, determining that the disease monitoring data x_i corresponding to the time point i is an abnormal value.

Embodiment four

FIG. 3 is a schematic structural diagram of a computing device provided in Embodiment four of the present disclosure. The computing device 1 may include: a memory 20, at least one processor 30, computer-readable instructions 40 stored in the memory 20 and executable on the at least one processor 30, for example, disease abnormal data detection program. The processor 30 executes the computer-readable instructions 40 to implement the steps in the embodiment of the disease abnormal data detection method, such as steps 101-105 shown in FIG. 1. Alternatively, the processor 30 executes the computer-readable instructions 40 to implement the functions of the modules/units in the foregoing device embodiments, such as the units 201-204 in FIG. 2.

Exemplarily, the computer-readable instructions 40 can be divided into one or more modules/units, and the one or more modules/units are stored in the memory 20 and executed by the at least one processor 30 to complete the present disclosure. The one or more modules/units can be a series of computer-readable instruction segments capable of performing specific functions, and the instruction segments are used to describe execution processes of the computer-readable instructions 40 in the computing device 1. For example, The computer-readable instruction 40 can be divided into the acquisition unit 201, the calculation unit 202, the determination unit 203, and the comparing unit 204 in FIG. 2. For the specific functions of each unit, refer to the Embodiment two.

The computing device 1 can be a computing device such as a desktop computer, a notebook, a palmtop computer, and a cloud server. Those skilled in the art will understand that the schematic diagram 3 is only an example of the computing device 1, and does not constitute a limitation on the computing device 1. Another computing device 1 may include more or fewer components than shown in the figures, or combine some components or have different components. For example, the computing device 1 may further include an input/output device, a network access device, a bus, and the like.

The at least one processor 30 can be a central processing unit (CPU), or can be another general-purpose processor, digital signal processor (DSPs), application specific integrated circuit (ASIC), Field-Programmable Gate Array (FPGA), another programmable logic device, discrete gate, transistor logic device, or discrete hardware component, etc.. The processor 30 can be a microprocessor or any conventional processor. The processor 30 is a control center of the computing device 1 and connects various parts of the entire computing device 1 by using various interfaces and lines.

The memory 20 can be configured to store the computer-readable instructions 40 and/or modules/units. The processor 30 may run or execute the computer-readable instructions and/or modules/units stored in the memory 20 and may call data stored in the memory 20 to implement various functions of the computing device 1. The memory 20 mainly includes a storage program area and a storage data area. The storage program area may store an operating system, an application program required for at least one function (such as a sound playback function, an image playback function, etc.), etc.. The storage data area may store data (such as audio data, a phone book, etc.) created according to use of the computing device 1. In addition, the memory 20 may include a high-speed random access memory, and may also include a non-transitory storage medium, such as a hard disk, an internal memory, a plug-in

hard disk, a smart media card (SMC), a secure digital (SD) Card, a flashcard, at least one disk storage device, a flash memory device, or other non-transitory solid-state storage device.

When the modules/units integrated into the computing device 1 are implemented in the form of software functional units and sold or used as independent products, they can be stored in a non-transitory readable storage medium. Based on this understanding, all or part of the processes in the methods of the above embodiments implemented by the present disclosure can also be completed by related hardware instructed by computer-readable instructions. The computer-readable instructions can be stored in a non-transitory readable storage medium. The computer-readable instructions, when executed by the processor, may implement the steps of the foregoing method embodiments. The computer-readable instructions include computer-readable instruction codes, and the computer-readable instruction codes can be in a source code form, an object code form, an executable file, or some intermediate form. The non-transitory readable storage medium can include any entity or device capable of carrying the computer-readable instruction code, a recording medium, a U disk, a mobile hard disk, a magnetic disk, an optical disk, a computer memory, a read-only memory (ROM).

In several embodiments provided in the preset application, it should be understood that the disclosed computing device and method can be implemented in other ways. For example, the embodiments of the devices described above are merely illustrative. For example, divisions of the units are only logical function divisions, and there can be other manners of division in actual implementation.

In addition, each functional unit in each embodiment of the present disclosure can be integrated into one processing unit, or can be physically present separately in each unit, or two or more units can be integrated into one unit. The above integrated unit can be implemented in a form of hardware or in a form of a software functional unit.

The present disclosure is not limited to the details of the above-described exemplary embodiments, and the present disclosure can be embodied in other specific forms without departing from the spirit or essential characteristics of the present disclosure. Therefore, the present embodiments are to be considered as illustrative and not restrictive, and the scope of the present disclosure is defined by the appended claims. All changes and variations in the meaning and scope of equivalent elements are included in the present disclosure. Any reference sign in the claims should not be construed as limiting the claim. Furthermore, the word "comprising" does not exclude other

units nor does the singular exclude the plural. A plurality of units or devices stated in the system claims may also be implemented by one unit or device through software or hardware. Words such as first and second are used to indicate names, but not in any particular order.

Finally, the above embodiments are only used to illustrate technical solutions of the present disclosure and are not to be taken as restrictions on the technical solutions. Although the present disclosure has been described in detail with reference to the above embodiments, those skilled in the art should understand that the technical solutions described in one embodiment can be modified, or some of the technical features can be equivalently substituted, and that these modifications or substitutions are not to detract from the essence of the technical solutions or from the scope of the technical solutions of the embodiments of the present disclosure.

CLAIMS

We claim:

1. A disease abnormal data detection method, the method comprising:

forming time series data X by acquiring disease monitoring data from time point 0 to time point t , $X = [x_0, x_1, x_2, \dots, x_t]$;

setting a time window size w , for each time point i from time point w to time point t , calculating a mean μ_i and a standard deviation σ_i of the disease monitoring data in a time window corresponding to the time point i , and the size of the time window corresponding to each time point being w , $i = w, w + 1, \dots, t$;

obtaining an mz list $[mz_w, mz_{w+1}, mz_{w+2}, \dots, mz_t]$ by calculating a moving-z-score value of disease monitoring data x_i corresponding to the time point i , according to the mean μ_i and the standard deviation σ_i of the disease monitoring data corresponding to each time point i , where $mz_i = \frac{|x_i - \mu_i|}{\sigma_i}$;

determining an abnormal value threshold of the time series data X according to the mz list;

determining whether a moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than the abnormal value threshold, and when the moving-z-score value of disease monitoring data x_i corresponding to time

point i is greater than the abnormal value threshold, determining the disease monitoring data x_i corresponding to the time point i is an abnormal value.

2. The disease abnormal data detection method of claim 1, wherein the method of determining an abnormal value threshold of the time series data X according to the mz list comprising:

calculating a preset quantile or a preset percentile of the mz list, and determining the preset quantile or the preset percentile as the abnormal value thresholds; or

calculate an average number of the mz list and determining the average number as the abnormal value threshold.

3. The disease abnormal data detection method of claim 1, further comprising:

obtaining abnormal values obtained by detecting the time series data by other disease abnormal data detection methods;

comparing the abnormal values obtained by the disease abnormal data detection method with the abnormal values obtained by the other disease abnormal data detection method;

obtaining a final abnormal value according to a comparison result of the abnormal value obtained by the disease abnormal data detection method and the abnormal values obtained by the other disease abnormal data detection method.

4. The disease abnormal data detection method of claim 1, further comprising:

obtaining regional abnormal values for regions by detecting the disease abnormal data of the regions, and obtaining regional subordinate hospitals abnormal values for regional subordinate hospitals by detecting the regional subordinate hospitals;

comparing the regional abnormal values with the regional subordinate hospitals abnormal values, and obtaining a final abnormal value according to a comparison result of the regional abnormal values and the regional subordinate hospitals abnormal values.

5. The disease abnormal data detection method of any one of claims 1 to 4, wherein the disease monitoring data comprising a number of consultations, a consultation rate, a number of infections, and an incidence rate of the disease.

6. The disease abnormal data detection method of any one of claims 1 to 4, wherein the method of acquiring disease monitoring data from time point 0 to time point t comprises:

establishing a disease monitoring network composed of multiple monitoring points in a preset area and acquiring disease monitoring data from the multiple monitoring points.

7. The disease abnormal data detection method according to claim 6, wherein the monitoring points comprising medical institutions, schools, childcare institutions, and pharmacies meeting a preset number or level.

8. A disease abnormal data detection device, the device comprising:

an acquisition unit, configured to form time series data X by acquiring disease monitoring data from time point 0 to time point t , $X = [x_0, x_1, x_2, \dots, x_t]$;

a calculation unit, configured to set a time window size w , for each time point i from time point w to time point t , calculate a mean μ_i and a standard deviation σ_i of the disease monitoring data in a time window corresponding to the time point i , and the size of the time window corresponding to each time point being w , $i = w, w + 1, \dots, t$;

a calculation unit, also configured to obtain an mz list $[mz_w, mz_{w+1}, mz_{w+2}, \dots, mz_t]$ by calculating a moving-z-score value of disease monitoring data x_i corresponding to the time point i , according to the mean μ_i and the standard deviation σ_i of the disease monitoring data corresponding to each time point i , where $mz_i = \frac{|x_i - \mu_i|}{\sigma_i}$;

a determination unit, configured to determine an abnormal value threshold of the time series data X according to the mz list;

a comparing unit, configured to determine whether a moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than the abnormal value threshold, and when the moving-z-score value of disease monitoring data x_i corresponding to time point i is greater than the abnormal value

threshold, determine the disease monitoring data x_i corresponding to the time point i is an abnormal value.

9. A computing device, the computing device comprising a memory and a processor, the memory stores at least one computer-readable instruction, and the processor executes the at least one computer-readable instruction to implement the following steps:

forming time series data X by acquiring disease monitoring data from time point 0 to time point t , $X = [x_0, x_1, x_2, \dots, x_t]$;

setting a time window size w , for each time point i from time point w to time point t , calculating a mean μ_i and a standard deviation σ_i of the disease monitoring data in a time window corresponding to the time point i , and the size of the time window corresponding to each time point being w , $i = w, w + 1, \dots, t$;

obtaining an mz list $[mz_w, mz_{w+1}, mz_{w+2}, \dots, mz_t]$ by calculating a moving-z-score value of disease monitoring data x_i corresponding to the time point i , according to the mean μ_i and the standard deviation σ_i of the disease monitoring data corresponding to each time point i , where $mz_i = \frac{|x_i - \mu_i|}{\sigma_i}$;

determining an abnormal value threshold of the time series data X according to the mz list;

determining whether a moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than the abnormal value threshold, and when the moving-z-score value of disease monitoring data x_i corresponding to time point i is greater than the abnormal value threshold, determining the disease monitoring data x_i corresponding to the time point i is an abnormal value.

10. The computing device of claim 9, wherein the method of determining an abnormal value threshold of the time series data X according to the mz list comprising:

calculating a preset quantile or a preset percentile of the mz list, and determining the preset quantile or the preset percentile as the abnormal value thresholds; or

calculate an average number of the mz list and determining the average number as the abnormal value threshold.

11. The computing device of claim 9, the processor is further used to execute the at least one computer-readable instruction to implement the following steps:

obtaining abnormal values obtained by detecting the time series data by other disease abnormal data detection methods;

comparing the abnormal values obtained by the disease abnormal data detection method with the abnormal values obtained by the other disease abnormal data detection method;

obtaining a final abnormal value according to a comparison result of the abnormal value obtained by the disease abnormal data detection method and the abnormal values obtained by the other disease abnormal data detection method.

12. The computing device of claim 9, the processor is further used to execute the at least one computer-readable instruction to implement the following steps:

obtaining regional abnormal values for regions by detecting the disease abnormal data of the regions, and obtaining regional subordinate hospitals abnormal values for regional subordinate hospitals by detecting the regional subordinate hospitals;

comparing the regional abnormal values with the regional subordinate hospitals abnormal values, and obtaining a final abnormal value according to a comparison result of the regional abnormal values and the regional subordinate hospitals abnormal values.

13. The computing device of any one of claims 9 to 12, wherein the disease monitoring data comprising a number of consultations, a consultation rate, a number of infections, and an incidence rate of the disease.

14. The computing device of any one of claims 9 to 12, wherein the method of acquiring disease monitoring data from time point 0 to time point t comprises:

establishing a disease monitoring network composed of multiple monitoring points in a preset area and acquiring disease monitoring data from the multiple monitoring points.

15. A non-transitory storage medium having stored thereon at least one computer-readable instructions that, when the at least one computer-readable instructions are executed by a processor to implement the following steps:

forming time series data X by acquiring disease monitoring data from time point 0 to time point t , $X = [x_0, x_1, x_2, \dots, x_t]$;

setting a time window size w , for each time point i from time point w to time point t , calculating a mean μ_i and a standard deviation σ_i of the disease monitoring data in a time window corresponding to the time point i , and the size of the time window corresponding to each time point being w , $i = w, w + 1, \dots, t$;

obtaining an mz list $[mz_w, mz_{w+1}, mz_{w+2}, \dots, mz_t]$ by calculating a moving-z-score value of disease monitoring data x_i corresponding to the time point i , according to the mean μ_i and the standard deviation σ_i of the disease monitoring data corresponding to each time point i , where $mz_i = \frac{|x_i - \mu_i|}{\sigma_i}$;

determining an abnormal value threshold of the time series data X according to the mz list;

determining whether a moving-z-score value of disease monitoring data x_i corresponding to the time point i is greater than the abnormal value threshold, and when the moving-z-score value of disease monitoring data x_i corresponding to time point i is greater than the abnormal value threshold, determining the disease monitoring data x_i corresponding to the time point i is an abnormal value.

16. The non-transitory storage medium of claim 15, wherein the method of determining an abnormal value threshold of the time series data X according to the mz list comprising:

calculating a preset quantile or a preset percentile of the mz list, and determining the preset quantile or the preset percentile as the abnormal value thresholds; or

calculate an average number of the mz list and determining the average number as the abnormal value threshold.

17. The non-transitory storage medium of claim 15, the at least one computer-readable instructions executed by the processor also to implement the following steps:

obtaining abnormal values obtained by detecting the time series data by other disease abnormal data detection methods;

comparing the abnormal values obtained by the disease abnormal data detection method with the abnormal values obtained by the other disease abnormal data detection method;

obtaining a final abnormal value according to a comparison result of the abnormal value obtained by the disease abnormal data detection method and the abnormal values obtained by the other disease abnormal data detection method.

18. The non-transitory storage medium of claim 15, the at least one computer-readable instructions executed by the processor also to implement the following steps:

obtaining regional abnormal values for regions by detecting the disease abnormal data of the regions, and obtaining regional subordinate hospitals abnormal values for regional subordinate hospitals by detecting the regional subordinate hospitals;

comparing the regional abnormal values with the regional subordinate hospitals abnormal values, and obtaining a final abnormal value according to a comparison result of the regional abnormal values and the regional subordinate hospitals abnormal values.

19. The non-transitory storage medium of any one of claims 15 to 18, wherein the disease monitoring data comprising a number of consultations, a consultation rate, a number of infections, and an incidence rate of the disease.

20. The non-transitory storage medium of any one of claims 15 to 18, wherein the method of acquiring disease monitoring data from time point 0 to time point t comprises:

establishing a disease monitoring network composed of multiple monitoring points in a preset area and acquiring disease monitoring data from the multiple monitoring points.