

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
13 February 2003 (13.02.2003)

PCT

(10) International Publication Number
WO 03/012436 A2

(51) International Patent Classification⁷: **G01N 33/48**

KING, James [US/US]; 42 Fawn Hill Lane, Monsey, NY 10952 (US). **BYRON, Karen** [US/US]; 3 Seminole Drive, Ringwood, NJ 07456 (US).

(21) International Application Number: PCT/US01/43352

(22) International Filing Date:
21 November 2001 (21.11.2001)

(74) Agents: **RODRICK, Richard**; Becton Dickinson and Company, 1 Becton Drive, Franklin Lakes, NJ 07417 et al. (US).

(25) Filing Language: English

(81) Designated States (*national*): CA, JP, US.

(26) Publication Language: English

(30) Priority Data:
60/308,987 31 July 2001 (31.07.2001) US

(84) Designated States (*regional*): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

(71) Applicant (*for all designated States except US*): **BECTON DICKINSON AND COMPANY** [US/US]; 1 Becton Drive, Franklin Lakes, NJ 07417 (US).

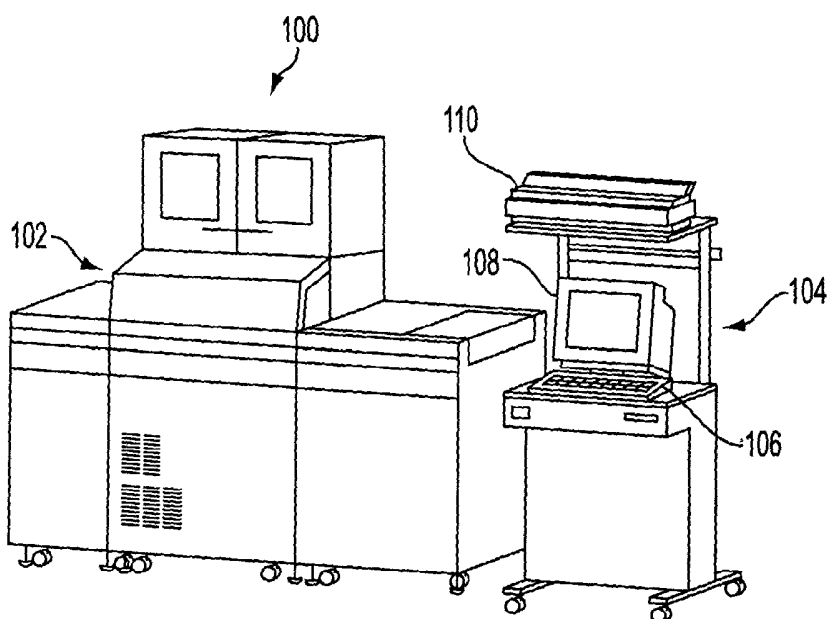
Published:
— *without international search report and to be republished upon receipt of that report*

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **BUSH, Valerie** [US/US]; 70 Lilly Road, Wanaque, NJ 07465-2412 (US).

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

(54) Title: METHOD AND SYSTEM FOR PREDICTING INITIAL ANALYTE VALUES IN STORED SAMPLES



(57) Abstract: A system and method are provided for estimating the initial level of an analyte in a sample based on an actual level observed, a time of storage of the sample, a storage temperature, and a type of container in which said sample is stored. Equations are generated for a particular analyte based on observations of actual analyte levels initially after a sample is drawn, as well as various times thereafter. The time of storage, the temperature of storage, and the type of container are observed along with actual analyte level in each observation. The equations generated are considered useful if the estimated initial values are more accurate than the actual observed final values.



WO 03/012436 A2

METHOD AND SYSTEM FOR PREDICTING INITIAL ANALYTE VALUES IN STORED SAMPLES

[0001] This application claims priority from U.S. Provisional Application Serial No. 60/308,987 filed July 31, 2001, the contents of which are expressly incorporated herein by reference in their entirety.

Field of the Invention

[0002] The present invention is related to biological and chemical sample analysis. In particular, but not by way of limitation, the present invention is related to a system and method for predicting initial values of common analytes in a blood

Background of the Invention

[0003] Many central laboratories receive specimens from remote collection stations, satellite laboratories, physician office laboratories (POLs), home health care nurses, and the like. Laboratory environments have changed such that more and more specimens are collected remotely and transported to core labs for analysis. In these situations, the specimens may not reach the testing laboratory until several hours or days after collection. Furthermore, the conditions of transport may vary widely. Additionally, stored specimens are often retrieved for add-on or repeat testing. In each of these situations, time and temperature can influence analyte stability, and hence the actual observed value of a specimen tested after transportation and/or storage.

[0004] Previous studies have been conducted to evaluate the stability of analytes. However, in those studies, either time or temperature was fixed. In the real world, the laboratory may have little control over the transport and storage conditions of specimens from remote locations.

[0005] Analyte values are used by physicians and medical professionals to make diagnoses of diseases and other health related problems. Proper diagnoses are dependent on accurate determinations of analyte values. Because analyte values can change over time, any determination made based on observed analyte values at a later time can be affected by the degree of analyte degradation since the sample was taken. Therefore, It would be desirable to have an accurate model for estimating the initial values of analytes in a blood sample, given knowledge of the transport and storage conditions. Such conditions include the type of container used to store the sample, the

temperature at which the sample was stored, and the time elapsed from when the sample was obtained to when it is actually tested.

[0006] Furthermore, it would be desirable to have testing equipment which incorporated such a model, to provide not only the actual observed value for a particular analyte, but also an accurate estimate of what the value was when the sample was taken, given knowledge of the transport and storage conditions, as discussed above.

Summary of the Invention

[0007] The above mentioned disadvantages are overcome and other advantages are realized by the present invention, which relates to a method for predicting an initial value of an analyte in a sample, particularly but not exclusively a blood sample. The method involves the steps of making a plurality of observations on a plurality of known samples, wherein each observation includes a plurality of variables associated with the sample. The variables include the actual analyte level at the time the sample was taken, the time since the sample was taken, the temperature at which the sample was stored, the type of tube or container in which the sample was stored, and the measured analyte level obtained after the storage time. The method also includes generating a model or equation which closely approximates the plurality of observations. The method also involves measuring the actual analyte level in an unknown sample, inputting the container type, storage time and storage temperature into the equation, and finally solving the equation to obtain the estimated initial analyte value.

[0008] The invention further comprises a method of predicting an initial analyte value given a sample that has been stored for a known time at a known temperature and in a known type of container. The method comprises the steps of determining the actual level of an analyte in a sample. Inputting data associated with the sample, including the storage time, storage temperature and container type, and applying the data to a predictive model or equation. The method further involves solving the equations to obtain an estimated initial analyte value.

[0009] The invention also comprises a system for estimating an initial value of an analyte in a sample. The system comprises a common chemistry analyzer which is adapted to analyze the actual level of one or more analytes in the sample. The system further comprises an estimator which is adapted to estimate the initial level analyte in the sample based on values including the actual level of the analyte measured by the analyzer, the storage time, the storage temperature, and the tube or container type.

[0010] The invention further comprises a method of generating a predictive model for predicting the initial value of an analyte in a sample. The samples are understood to preferably be blood samples, but the invention is not limited to blood samples. The method comprises the steps of collecting a plurality of samples, testing each sample for an initial value of an analyte, storing at least one sample at each of a plurality of storage temperatures, testing each sample for subsequent levels of the analyte after known time intervals, analyzing the data collected based on a polynomial regression model, and generating a predictive model or equation based on the results.

[0011] The invention further relates to a computer-readable medium of instructions which are adapted to control a system to generate a predicted initial

analyte value. A first set of instructions are adapted to control the system to collect a plurality of data associated with a sample, the data comprising an actual analyte value, a time of storage, a temperature of storage, and a type of container in which the sample was stored. A second set of instructions are adapted to control the system to apply the data to a predictive model and to calculate an estimated initial analyte value. A third set of instructions are adapted to control the system to output the estimated initial analyte value.

[0012] The computer-readable medium of instructions can be further adapted to generate the predictive model. In this case, the computer-readable medium of instructions further includes a fourth set of instructions adapted to control the system to collect a plurality of data associated with a known set of samples. Each sample is associated with data including an actual initial analyte value, an actual subsequent analyte value, a time of storage, a temperature of storage, and a type of container. A fifth set of instructions is adapted to control the system to generate the predictive model from the data associated with the known set of samples.

Brief Description of the Drawings

[0013] The invention will be further understood with reference to the attached drawing figures, in which:

[0014] Figure 1 illustrates a system according to an embodiment of the present invention;

[0015] Figure 2 is a flowchart depicting a method of producing a predictive model in accordance with an embodiment of the invention;

- 6 -

[0016] Figure 2(a) is a plot of analyte value vs. time for various storage temperatures, in accordance with an embodiment of the present invention;

[0017] Figures 3(a)-(g) are histograms comparing the differences between predicted initial results and actual initial results to the differences between aged actual results and actual initial results;

[0018] Figure 4 is a flowchart depicting a method of validating a predictive model as used in an embodiment of the present invention.

[0019] Figure 5 is a flowchart depicting a method of predicting initial analyte values using a predictive model in accordance with an embodiment of the present invention;

[0020] Throughout the drawing figures, it will be understood that like numerals refer to like structures or steps.

Detailed Description of the Invention

[0021] The following description of the presently contemplated best mode of practicing the invention is not to be taken in a limiting sense, but is provided merely for the purpose of describing the general principles of the invention. The scope of the invention should be determined by reference to the claims.

[0022] Figure 1 shows a system according to an embodiment of the present invention. The system 100 includes blood analyzing equipment 102, as well as a computer 104. The computer has input devices, such as the keyboard 106 shown, and output devices, such as the monitor 108 and printer 110 shown. It will be understood that a wide variety of input devices could be employed, as is known to those of skill in

- 7 -

the art. For instance, the input devices could include bar code readers, RF tag readers, optical character recognition (OCR) equipment, and any other type of input device which could provide data to the computer 104. Similarly, the output devices shown include a monitor 108 and printer 110, and it should be understood that these are merely shown as examples, but that a wide variety of output devices are contemplated to be within the scope of the invention. Some input devices, such as a bar code reader, are particularly advantageous for their ability to input data associated with particular blood samples. Blood samples could be provided with bar codes containing data related to the sample, such as the actual time and date the sample was drawn, and the type of container used to hold the sample, and the temperature at which the sample was stored.

[0023] Of course it should be understood that while a system according to an embodiment of the invention is described above as including a computer, the invention is not intended to be limited to a computer system. Simplified implementations of the invention could included a spreadsheet with a set of formulas designed to calculate estimated initial analyte values from a set of input values, such as temperature and time of storage. Similarly, the invention could comprise a web page having input fields, and an application running on the web page's server which calculated estimated initial values of analytes from information input by a user navigating to the web page. The above examples are provided merely as illustrations, and are meant in no way to limit the scope of the invention.

[0024] As will be described in further detail below, the invention may be understood to comprise two phases. The first phase involves generating and validating

predictive models for particular blood analytes, which accurately and reliably predict the initial value of an analyte in a blood sample, given the actual value measured at a later time, along with data including the storage time, storage temperature, and container type. The second phase involves employing the model on a new blood sample which has been stored for a known time at a known temperature in a known container, in order to obtain an accurate and reliable estimate of what a blood analyte's value was when the sample was first obtained.

Generating Predictive Model

[0025] Figure 2 is a flowchart illustrating the process of producing a predictive model, or equation, for a given analyte. The first step 200 in producing an equation for a particular blood analyte is to design and conduct a study that incorporates the significant variables, such as tube type, storage time, and storage temperature. Examples of a study will be described in greater detail below. Next, after the data is collected, the next step 202 is to generate an equation as a function of all of the inputs. This is done through statistical analysis of the data. A "best fit" equation is produced based on the observations made during the study. Next, Statistical tests are used to reduce the terms in the equation at step 204. Next, the equation is inverted at step 206 to solve for the initial analyte value. Finally, at step 208, the coefficients of the equation representing the predictive model are outputted. More detailed information on regression analysis can be found in Dunn, et al., Basic Statistics: A Primer for the Biomedical Sciences, Third Edition, which is incorporated herein in its entirety.

[0026] The following is a description of a study conducted to generate predictive models, in accordance with the present invention. Twenty one common chemistry blood analytes were studied, along with two hormones. The analytes are listed in Table 1 below:

Table 1. Analytes Tested/Validated	
Sodium	Total Protein
Potassium	Albumin
Chloride	Total Bilirubin
Bicarbonate	Alk. Phos.
Calcium	AST
Phosphate	ALT
Magnesium	GGT
Urea	Iron
Creatinine	Cholesterol
Uric Acid	Triglyceride
Glucose	Free T4
Estradiol	

[0027] Blood was collected from twenty healthy donors, and placed into three different types of tubes in randomized order. Tubes were stored at storage temperatures of -20°C , 4°C , 25°C and 40°C . Actual analyte values were measured at times of 0, 8, 24, 48, and 168 hours after collection. From the healthy donors, 171 specimens were collected into the three different types of tubes, with 57 samples per tube type. Analysis of the 21 common chemistry analytes was performed on an Olympus 5000. The two hormones were analyzed on an Abbott Imx.

[0028] Once the data was collected, a polynomial regression model was fit to the data. The model included the significant factors allowed by the study. These included quadratic terms, such as Temp^2 and Time^2 , as well as interaction terms, such as $\text{Temp}*\text{Time}$.

[0029] The following set of equations were generated for a particular analyte, Calcium. One equation is presented for each tube type:

Calcium	Plain Tube	$Ca = 9.47855 - 0.00698 (\text{Time}) + 0.00024 (\text{Temp}) + 0.00004 (\text{Temp})^2$
	SST Tube	$Ca = 9.39084 - 0.00698 (\text{Time}) + 0.00482 (\text{Temp}) + 0.00004 (\text{Temp})^2$
	Eval Tube	$Ca = 9.27934 - 0.00698 (\text{Time}) + 0.00914 (\text{Temp}) + 0.00004 (\text{Temp})^2$

Of course it will be understood that the above equations could be presented in other formats. For instance, a single equation could be presented which takes tube type into account as a variable. Furthermore, an equation or set of equations could be presented which do not include temperature as a variable, but rather assume a certain temperature. In other words, a set of equations could be presented, one equation per temperature. Of course any variety of equations could be presented, depending on the needs of a particular situation. The above representative equation is not meant to be limited, but rather is presented merely as illustrative of the invention. A wide variety of equations are contemplated to be within the scope of the invention.

[0030] Figure 2(a) illustrates a plot of analyte values vs. time for several temperature values, for the analyte CO₂. In the plot, each line is representative of analyte values over time for a particular storage temperature. The plots are made based on equations generated as described above.

Validation Of Predictive Model

[0031] The models were validated with independent patient data sets. Initial results for each patient were predicted based on the patient result observed at a latter time, the storage time, the storage temperature, and the type of tube used for sample

collection. The true initial results were known for the validation data, and the differences between the model-based predictions and the true initial results were calculated. The percent values occurring within clinical acceptance limits of the initial value (time = 0) were calculated for repeat observed values and the predicted values.

[0032] To evaluate the models, the differences between the model predictions and the true initial results were compared to the differences between the aged sample results and the true initial results. This mimics the real situation where the aged sample result substitutes for the initial result. Vertically stacked histograms, shown in Figs. 3(a)-(g), are used to compare the two sets of differences. The histograms shown in Figs. 3(a)-(g) represent both precision and accuracy of the estimate for a typical patient. The donor's "true" value is the midpoint of the interval shown on the horizontal axis. The histograms in the top row of each of Figures 3(a)-(g) represent the differences between the aged observed sample results and the true observed initial results. The histograms in the bottom rows represent the differences between the model-adjusted predictions of initial value and the true observed initial results.

[0033] The process of validating predictive models is illustrated in Figure 4. First, at step 300, a set of samples are taken, and data are collected. The set of samples taken for validation is preferably different from the set used to generate the predictive models. In the validation set of data, the initial analyte value, the storage time, temperature, and storage tube type, as well as the final analyte value are all known. At step 302, the predictive model is used to estimate the initial values of analytes in the validation set. At step 304, the estimated initial value is compared to the actual observed final value. Finally, at step 306, the equation is determined to be useful if the

estimated (predicted) initial values are more accurate and more precise than the actual observed final values.

Using the Predictive Model

[0034] The flowchart of Figure 5 illustrates the process of using the predictive models to determine an unknown initial analyte value from a sample which has been stored under known conditions, including known time, temperature, and tube type. At step 400, the sample is received, including data related to the storage history of the sample. Analysis is performed on the sample to determine actual present analyte values. The temperature and time of storage, and the tube type are checked at step 402 to determine if they are within study limits. If they are not, the predictive model cannot reliably be used (step 404). If the values are within study limits, then at step 406, the actual analysis result, along with data pertaining to the tube type, storage time, and storage temperature are input to the predictive model. The equation is calculated, and at step 408, the estimated initial analyte value is output.

[0035] While the invention herein disclosed has been described by means of specific embodiments and applications thereof, numerous modifications and variations could be made thereto by those skilled in the art without departing from the scope of the invention set forth in the claims.

WHAT IS CLAIMED IS:

1. A method of predicting an initial value of an analyte in a sample, said method comprising the steps of:
 - making a plurality of observations on a plurality of samples, wherein each observation includes a plurality of variables associated with said sample, said variables comprising an actual initial analyte level, an actual subsequent analyte level, and at least one of the factors taken from the group consisting of a time since the sample was taken, a temperature at which said sample was stored, and a type of container in which said sample was stored,
 - generating an equation which approximates said plurality of observations,
 - analyzing a sample analyte value, said sample having associated therewith a container type, a storage time, and a storage temperature,
 - inputting said container type, said storage time, said storage temperature, and said analyte value into said equation, and
 - solving said equation to obtain an estimated initial analyte value.
2. The method of claim 1, further comprising the step of solving said equation for initial analyte value as a function of container type, storage time, and storage temperature.
3. The method of claim 1, wherein said sample is a blood sample.
4. The method of claim 1, wherein the plurality of variables associated with said sample comprises comprising an actual initial analyte level, an actual subsequent analyte level, a time since the sample was taken, and a temperature at which said sample was stored.
5. The method of claim 1, wherein the plurality of variables associated with said sample comprises comprising an actual initial analyte level, an actual subsequent

analyte level, a time since the sample was taken, and a type of container in which said sample was stored.

6. The method of claim 1, wherein the plurality of variables associated with said sample comprises comprising an actual initial analyte level, an actual subsequent analyte level, a temperature at which the sample was stored, and a type of container in which said sample was stored.
7. The method of claim 1, wherein the plurality of variables associated with said sample comprises comprising an actual initial analyte level, an actual subsequent analyte level, a time since the sample was taken, a temperature at which said sample was stored, and a type of container in which said sample was stored.
8. A method of predicting an initial value of an analyte in a sample, said method comprising the steps of:
 - determining a level of an analyte in a sample, said sample having associated therewith at least one of the factors out of the group consisting of a container type, a storage time, and a storage temperature,
 - inputting said factors and said analyte level into an equation, and
 - solving said equation to obtain an estimated initial analyte value.
9. The method of claim 8, wherein said equation represents analyte level as a function of storage time.
10. The method of claim 8, wherein said equation represents analyte level as a function of storage temperature.
11. The method of claim 8, wherein said equation represents analyte level as a function of container type.

12. The method of claim 8, wherein said equation represents analyte level as a function of storage temperature and storage time.
13. The method of claim 8, wherein said equation represents analyte level as a function of storage temperature and container type.
14. The method of claim 8, wherein said equation represents analyte level as a function of storage time and container type.
15. The method of claim 8, wherein said equation represents analyte level as a function of storage temperature, storage time, and container type.
16. The method of claim 8, wherein said sample is a blood sample.
17. A system for estimating an initial value of an analyte in a sample, said system comprising:
 - a chemistry analyzer adapted to analyze the actual level of an analyte in a sample,
 - a estimator adapted to estimate said initial level of said analyte based on a plurality of values, said values including said actual level of said analyte, a storage time, a storage temperature, and a container type, and
 - an output adapted to present said estimated initial value.
18. The system of claim 17 wherein said estimator estimates said initial level using an equation associated with said analyte, said equation adapted to solve for initial analyte level as a function of actual analyte level, storage time, storage temperature, and container type.
19. The system of claim 18, wherein said equation contains a quadratic term for time, a quadratic term for temperature, and a mixed term for time and temperature.

20. The system of claim 17, wherein said sample is a blood sample.
21. A method of generating a predictive model for predicting the initial value of an analyte in a sample, the method comprising the steps of:
- collecting a plurality of samples,
 - testing each sample for an initial value of an analyte,
 - storing at least one sample at a known storage temperature,
 - testing each sample for subsequent values of said analyte after at least one additional time interval,
 - analyzing data based on said tests according to a polynomial regression model,
 - generating a predictive model based on said analysis.
22. The method of claim 21, further comprising the steps of:
- validating the predictive model by comparing the true initial values to a set of model predicted values.
23. The method of claim 22, wherein the step of validating further includes:
- accepting the model of the set of model predicted values is more accurate than the set of subsequent values.
24. The method of claim 21 wherein said sample is a blood sample.
25. The method of claim 21, wherein the step of storing at least one sample further comprises storing said samples at a plurality of known temperatures.
26. A computer-readable medium of instructions, adapted to control a system to generate a predicted initial analyte value, comprising:
- a first set of instructions, adapted to control said system to collect a plurality of data associated with a sample, said data comprising an actual analyte value, and

at least one factor from the group consisting of a time of storage, a temperature of storage, and a container type;

a second set of instructions, adapted to control said system to apply said data to a predictive model and to calculate an estimated initial analyte value; and

a third set of instructions, adapted to control said system to output said estimated initial analyte value.

27. A computer-readable medium of instructions, as in claim 26, further adapted to generate said predictive model, and further comprising:

a fourth set of instructions, adapted to control said system to collect a plurality of data associated with a known set of samples, said known samples each having data associated therewith comprising an actual initial analyte value, an actual subsequent analyte value, and at least one factor from the group consisting of a time of storage, a temperature of storage, and a container type;

a fifth set of instructions, adapted to control said system to generate said predictive model from said data associated with said known set of samples.

28. A computer-readable medium of instructions, as in claim 27, further comprising a sixth set of instructions, adapted to control said system to validate the predictive model by comparing the true initial values to a set of model predicted values.

29. A computer-readable medium of instructions, as in claim 28, further comprising a seventh set of instructions, adapted to control said system so accept the model if the set of model predicted values is more accurate than the set of subsequent values.

30. A computer-readable medium of instructions, as in claim 29, wherein the sample is a blood sample.

1/7

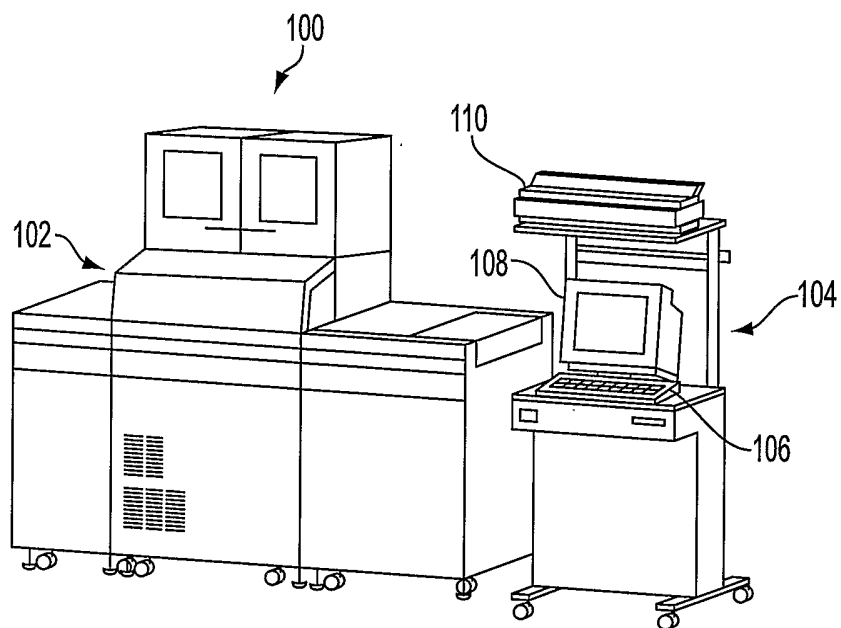


FIG. 1

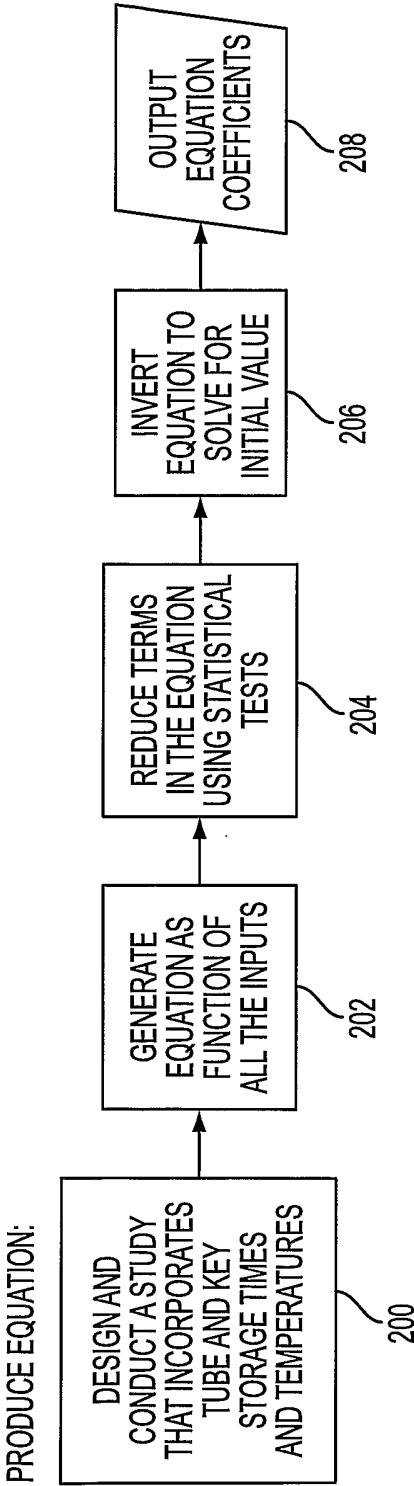


FIG. 2

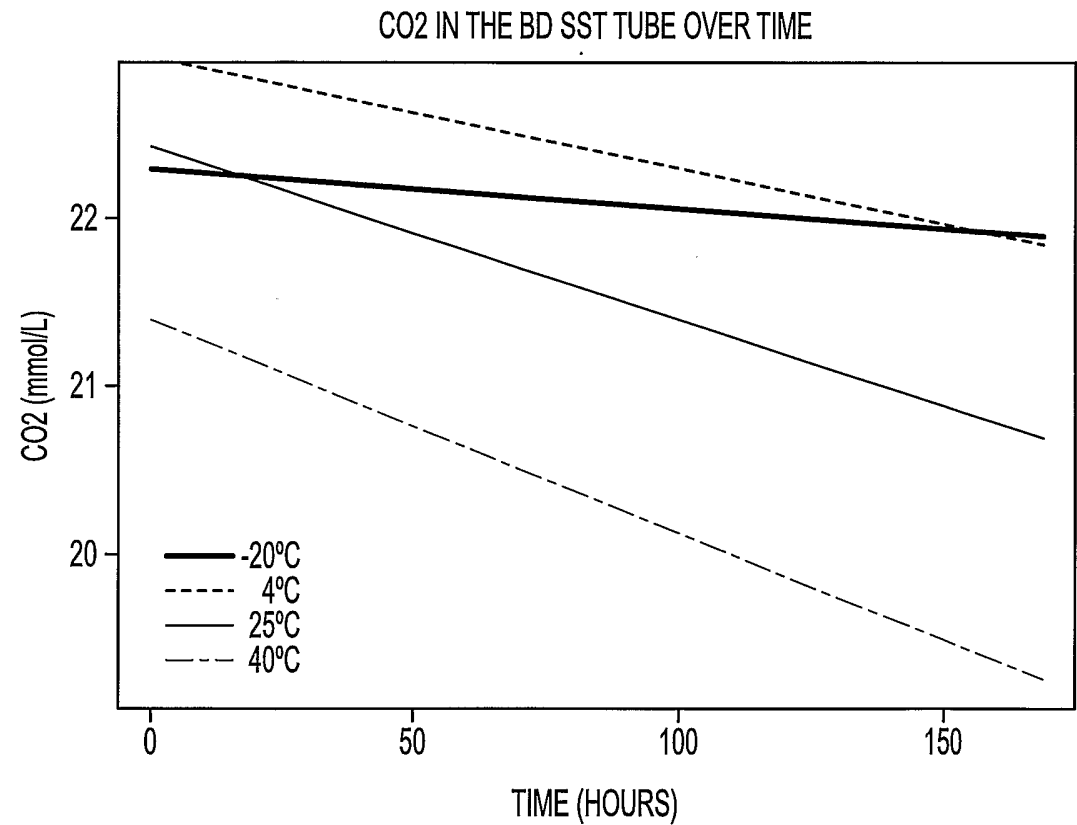


FIG. 2(a)

4/7

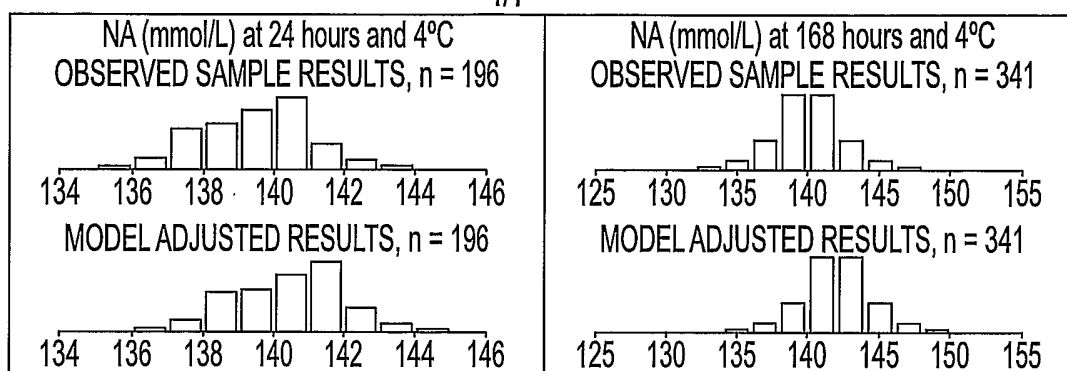


FIG. 3(a)

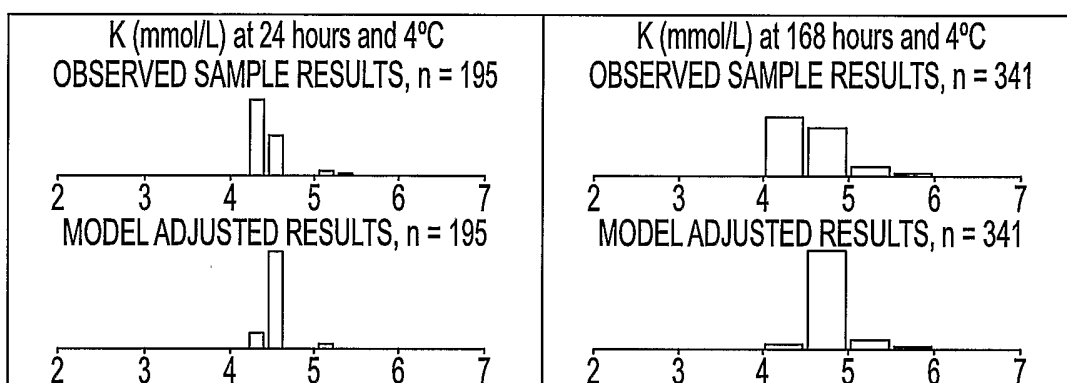


FIG. 3(b)

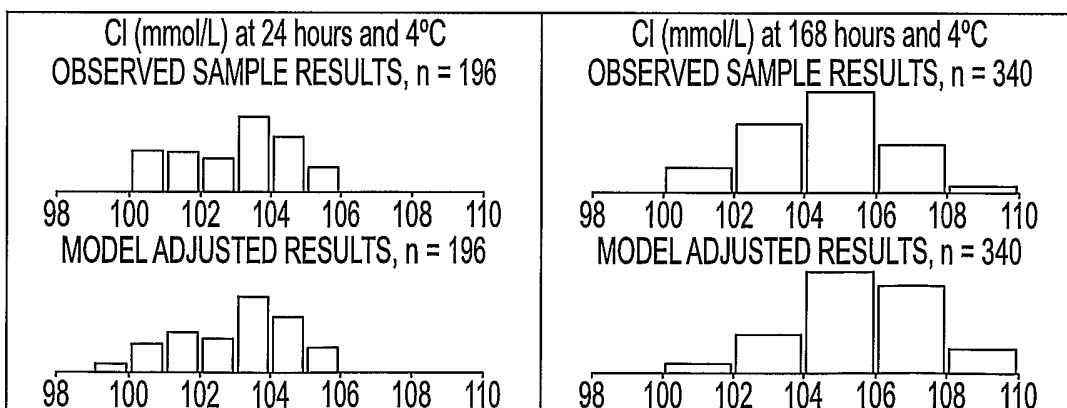


FIG. 3(c)

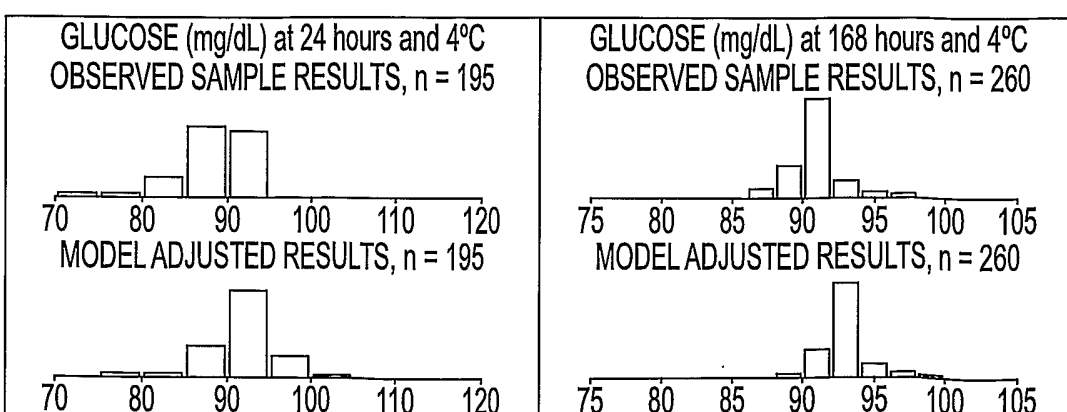


FIG. 3(d)

5/7

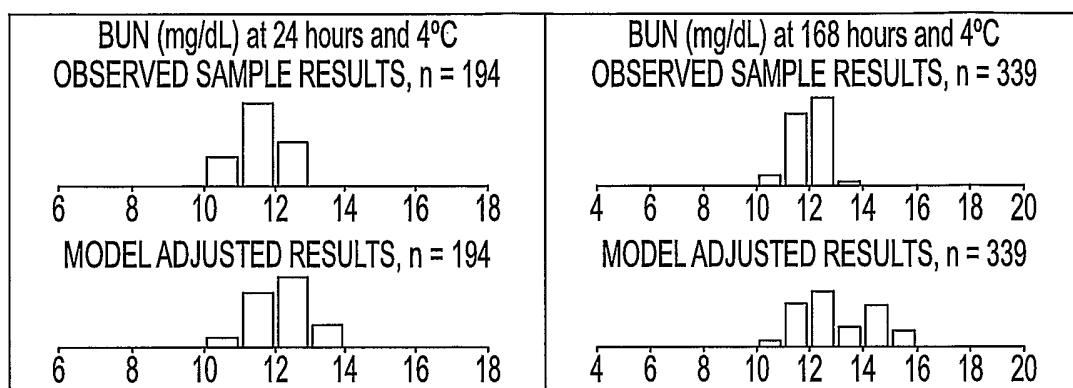


FIG. 3(e)

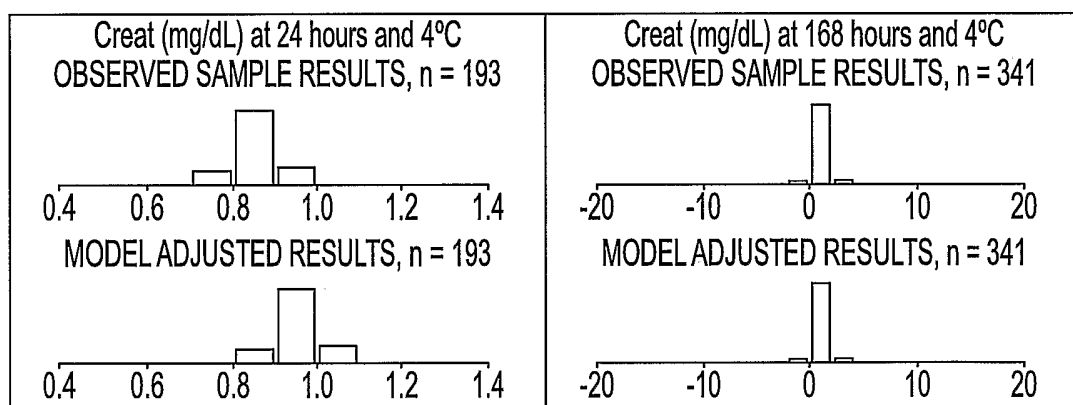


FIG. 3(f)

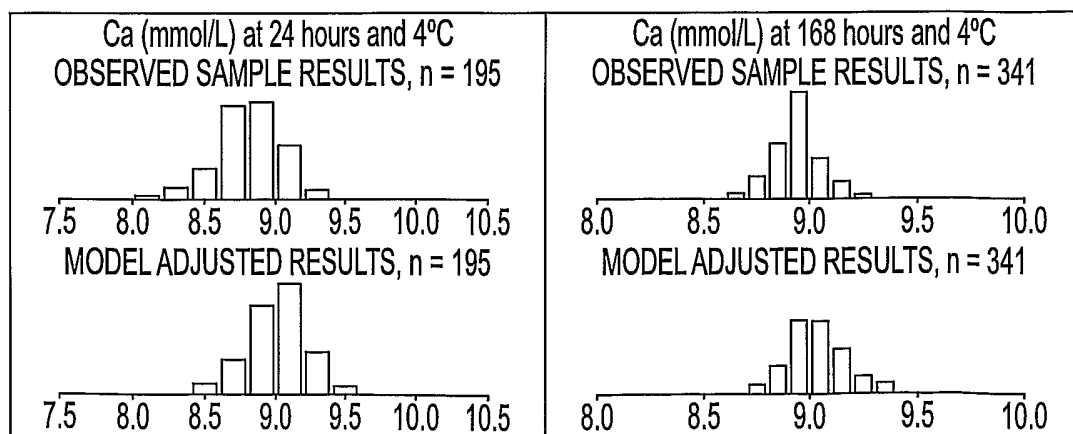


FIG. 3(g)

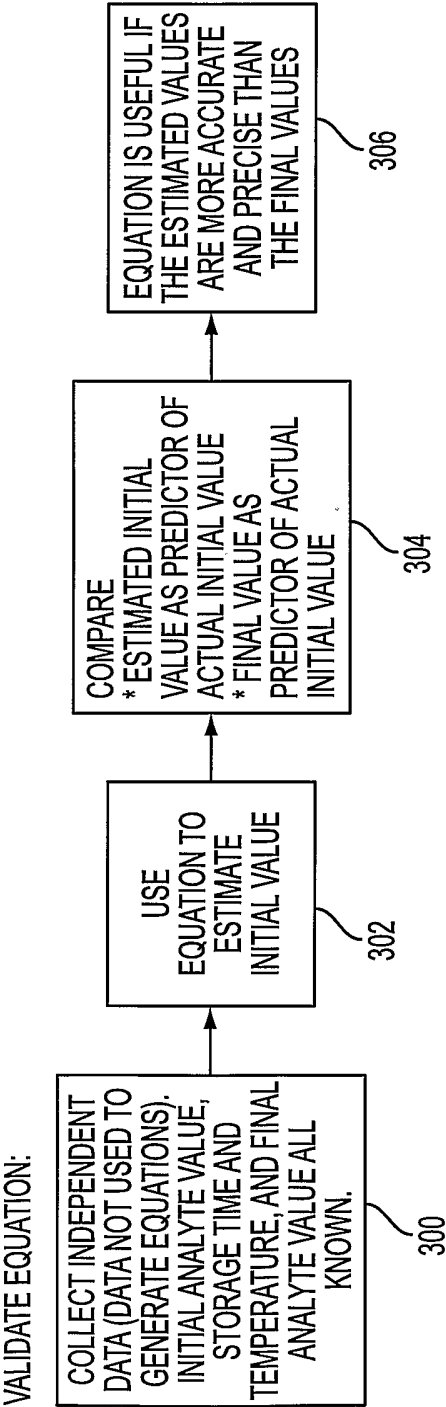


FIG. 4

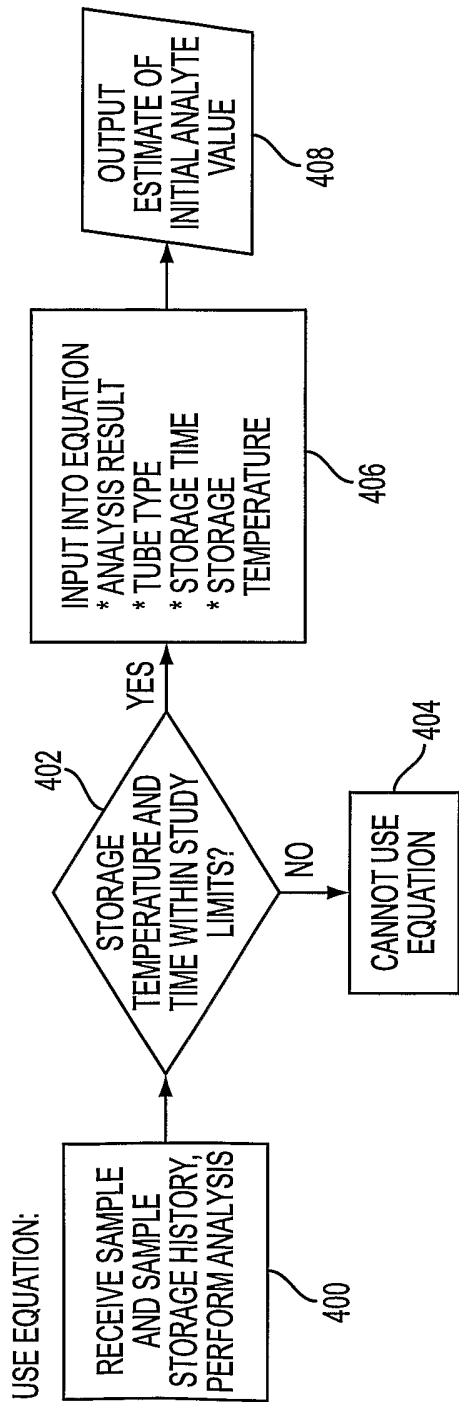


FIG. 5