Variance Function Program

Version 19.0

(for Windows 8.2 and later)

February 2023

W.A. Sadler
71B Middleton Road
Christchurch 8041
New Zealand

Ph: +64 3 343 3808
e-mail: bill.sadler@xtra.co.nz

(formerly at Nuclear Medicine Department, Christchurch Hospital)

Contents Page

Variance Function Program 19.0

The variance function (the relationship between variance and the mean) has several applications in statistical analysis of medical laboratory data (eg. Ref. *1*), but the single most important use is probably the construction of (im)precision profile plots (*2*). This program (VFP.exe) was initially aimed at immunoassay data which exhibit relatively extreme variance relationships. It was based around the "standard" 3-parameter variance function,

$$\sigma^2 = (\beta_1 + \beta_2 U)^J$$

where σ^2 denotes variance, U denotes the mean, and β_1, β_2 and J are the parameters (*3, 4*). It has since been extended to incorporate additional variance functions including simpler functions that might be better suited to the less extreme variance relationships exhibited by laboratory tests other than immunoassays (see **The Variance Function** on page 5). Following pages give a general description of the program followed by details of how the variance function might be usefully employed for statistical analysis. Appendices A, B and C give program delivery, installation and removal information, respectively. Appendices D and E describe the example files that are installed with the program. Appendix F gives a brief history of the program. Appendix G lists changes in the version 18.0 to version 19.0 upgrade and Appendix H describes bug fixes in the upgrade to version 19.0.

Windows Vista through Windows 11 Issues

Whenever VFP.exe starts up it checks the Windows registry and, if necessary, updates (or attempts to update) the association between itself and the files it creates (.VFD, VFS, VFR and .VFG; see **Program Structure** below). This means that the program can be started by just double clicking any of its files in Windows Explorer or by dragging a file onto the program icon or shortcut. More than one version of VFP.exe could be used on the same machine and/or the location of VFP.exe could be changed, because the association between the current instance and the program files is continuously updated. However, under Windows Vista and later, the User Account Control (UAC) prevents programs writing to the Windows registry. Accordingly, the association between VFP.exe and its program files is now defined during program installation and since this requires the location of VFP.exe to be known in advance, it is now automatically installed into directory C:\VFP (previously C:\VFP was suggested as the location but users could specify an alternative location if they wished).

If file associations need updating under Windows Vista or later then, prior to starting the program, right click the program icon or shortcut and select the "Run as Administrator" option (this assumes that you have administrator privileges).

Program Help

The program has an extensive on-line Help file (VFPHelp.chm). A brief popup description of all controls on all dialogs can be obtained by pressing F1 when the control has focus, or by clicking the right mouse button over the control (provided it is enabled), or by using the **?** icon at the top right of the dialog. All but the simplest dialogs have a Help button which gives access to detailed information about the dialog. Print copies of help topics if required. Pause the mouse over any speed button on the program toolbars for a transient popup description of the button's purpose; a fuller description appears on the lower left status bar.

When working with a program spreadsheet or with any of the program graphs, just click the right mouse button anywhere on the display area for immediate access to editing or design options. Drag-and-drop can be used in many contexts including positioning graph labels and ID elements, positioning and resizing inset plots. It can be used in all situations where a selection is made by transferring items from one list to another (eg. selecting variance functions, data point styles, graph colours, etc.) and also to change the ordering of selected items.

On-line help topic Contents > Overview contains general information about the program plus technical information about variance function estimation, variance stabilisation, Limit of Blank (LoB) and Limit of Detection (LoD) calculations, regression and Bland-Altman analysis. Help topic Contents > Automation gives complete technical information about implementing auto-processing. Other topics refer mainly to the fine detail of using the program, ie. the information produced when Help buttons are clicked.

Gestures

All graphs are accompanied by design dialogs which give considerable control over graph content and appearance. As from program version 12.0, the simplest manipulations can also be accomplished by simple manual gestures whenever a monitor has touch-screen capability (or by using the mouse, with left button depressed, if not). For example, toggle between linear and logarithmic X-axis scales (when both are permitted) using an L-shaped gesture where the vertical stroke cuts the X-axis. Similarly, toggle between linear and logarithmic Y-axis scales (when both are permitted) using an L-shaped gesture where the horizontal stroke cuts the Y-axis. Increase or decrease data point size (when relevant) using a vertical upward or downward stroke within the rectangle formed by the graph axes. Gestures are consistent across all graph types. Consult on-line help topic Contents > Overview > Gestures for a complete list of available gestures. Click Help buttons on individual graph dialogs to access information on which gestures apply to that graph type.

Program Structure

The program is written in the Object Pascal language using Embarcadero's RAD Studio 11.2 development environment. It consists of two modules; a *Data* module and a *Graphs* module. Click the right-most speed button on the program toolbar to switch between the two. A toolbar speed button, or main menu item Options > General Program Options, allows program global options to be selected and these include specifying which module should appear when the program is started and automatic reopening of the most recently open file when switching back and forth between the modules.

Data Module

The Data module allows creation and maintenance of three specialised spreadsheets;

(1) Data file (saved to disk with a .VFD extension). Stores runs of raw replicated measurements (max: 200000 observations).

(2) SD file (saved to disk with a .VFS extension). This spreadsheet has precisely three rows for entry of N, mean and SD values (max: 100000 sets). Intended for situations where only summarised precision data are available (eg. precision data published in the literature).

(3) Regression file (saved to disk with a .VFR extension). This spreadsheet has precisely two columns for entry of paired X, Y values (max: 100000 pairs).

The program can directly read any of these data categories from database tables (requires an ODBC link) or from Text (Ascii) files. Data can be copy/pasted from external applications into any of the program spreadsheets and can be copy/pasted from any of the program spreadsheets into external applications, or exported in Text file format.

Use the data module in the first instance to enter or import data and to view and output data summaries. Speed buttons on the program toolbar give immediate access to variance function estimation, regression analysis, LoB, LoD estimation, Bland-Altman calculations and variance stabilisation transformations. View and output calculation summaries and associated graphs (examples of which appear later in this document). Speed buttons also give access to histogram plots of sets of replicates, error plots, biological variation plots and to a dialog for organising the export of variance function parameters (sufficient to reproduce variance function plots in external applications). Descriptive notes about the data in Data, SD and Regression files can also be entered or viewed at any time.

Apart from clicking speed buttons, all the foregoing activities can also be accessed via the main menu items Summaries or Calculations & Plots, or via a popup menu invoked by clicking the right mouse button anywhere on the spreadsheet (the latter also gives immediate access to data editing help). Data, SD and Regression files saved to disk contain the data, notes, all details of variance function estimation, LoB, LoD, regression and Bland-Altman calculations, plus all design settings applied to variance function quick plots and regression and Bland-Altman plots.

Twenty-five example Data files (.VFD), one SD file (.VFS) and three Regression files (.VFR) are installed with the program (see Appendix D on page 16). File CRPEval.VFD, kindly donated by Jill Tate, Royal Brisbane and Women's Hospital, Brisbane, Australia, provides an example of a variance turning point near zero (Fig. 4 on page 6). Peter Watkinson, The John Radcliffe Hospital, Oxford, UK kindly provided Glucose.VFD, a set of arterial glucose data which illustrates the use of one of the simpler variance functions (Fig. 6 on page 7). The other files contain old evaluation and internal QC data from this laboratory and are intended to illustrate a few points about data or the program (refer to the notes stored in each file). They can also be used to illustrate data summaries and Data module graphical output including the design options that are available. Some example blank specimen data are provided in read-only MS Excel file, Blanks_Data.xls, which are intended to illustrate the importing and processing of blank specimen data (see also Limit of Blank (LoB) and Limit of Detection (LoD) on page 14).

Graphs Module

Use the Graphs module to construct and maintain variance function graphs. In contrast to Data module quick plots, which display a single variance function, up to 12 variance functions can be assigned to a graph. In essence, a graph is just a display container for the data and variance functions contained in one or more Data or SD files. The Graphs module gives greater control over the ranges of the X, Y axes and automates production of blow-up views of the origin region (often of special interest). Descriptive notes about a graph (eg. a figure legend) can be entered or viewed at any time. Graphs saved to disk have a .VFG extension. Graphs can be saved in two distinct forms; Unlocked and Locked.

Unlocked graphs refer to the variance functions assigned to them by storing only the disk location and names of the Data or SD files. When an Unlocked graph is reopened, the specified Data or SD files are read from scratch so that the graph will immediately reflect any updates to the data and variance functions. Save graphs in an Unlocked state when you want automatic updating as data accumulate (eg. internal QC data). Variance functions can be added to or removed from Unlocked graphs at any time. However, if a Data or SD file is deleted or moved the variance function will disappear from any Unlocked graphs to which it may have been assigned.

Locked graphs store not only the names of the Data or SD files but all data and variance function details contained in them. They are completely self contained. Save graphs in a Locked state when you want to guarantee being able to reproduce them regardless of changes to, or the availability of, the underlying Data or SD files (eg. graphs prepared for a report or for publication). Locked graphs cannot be unlocked. Variance functions in a Locked graph can be freely removed from the graph display and reselected, but you are limited to the specific data and variance functions that were locked into the graph.

Data Type

A key property in the Variance Function Program. It is compulsory for all data files in the sense that the Data Type (max: 40 characters) must be defined before the program allows estimation of the variance function, plotting, or files to be saved to disk. The Data Type is intended to link all files that contain data for a particular analyte, regardless of the instrument or assay method used to produce the data. Use file names, variance function labels or descriptive file headers to distinguish between instruments, assay types or other sub-categories into which data for a particular analyte might be classified. Apart from identifying the data, the Data Type is also used as the *default* X-axis label on all program graphs. A Data Type definition should therefore be thought of as the label that would be used if the data were plotted. eg. if the data are replicated measurements of the pituitary hormone thyrotropin (TSH) then graphs would typically be labelled;

Thyrotropin (mU/L) or **TSH (mU/L)**

Either of these would be suitable as the Data Type definition for thyrotropin data. It is probably best to use full names for the Data Type definition. If necessary, X-axis labels can be shortened to an abbreviated form, or otherwise modified to identify an instrument or assay method or experimental context, etc. Modifying an X-axis label does not affect the Data Type assigned to the data. Likewise, graph files in the program Graphs module must also be assigned a Data Type when created. Only variance functions with the same Data Type can be assigned to the graph. Putting that another way, any combinations of variance functions for a particular analyte can be assigned to a graph for comparison purposes, provided they have all been assigned the same Data Type. The Data Type assigned to X, Y pairs in a Regression file is used to limit the variance functions that can be specified as regression weighting functions. Likewise, when specifying a variance function to normalise X, Y pair differences in Bland-Altman analysis the choice is limited to those with the same Data Type.

A Data Type definition should only have to be entered once. During startup the program scans all Data, SD, Regression and Graph files plus auto-processing protocols, and adds previously defined Data Types to alphabetical drop down lists in dialogs which require the entry or selection of a Data Type. Browse the list to check whether a Data Type has already been defined. If so, select it to ensure the same definition is used for a particular analyte. Drop down lists should contain at least seven Data Type definitions when the program is run for first time (extracted from the example files and auto-processing protocols that are installed with the program). Refer to on-line help topic Contents > Overview > Data Type for more information.

Automation

Following program installation, interested users should consult on-line help topic Contents > Automation > Automated Data Processing – Overview. Print a copy of the overview topic as a reference. If you wish to use the example files that are provided to illustrate auto-processing (see Appendix E on page 17) then the program must be located in directory C:\VFP. If not, example Pointers and Protocols (fully explained in the overview help topic) will require modification and the example graph (.VFG) files will not "find" their assigned variance functions, which rather defeats the intention of providing an uncomplicated introduction to automation. Example auto-processing files include 22 raw input data files (text files) which contain internal QC data from this laboratory.

Program Data Area

Figure 1 shows the directory structure created (if necessary) by the program during installation, where Home (C:\VFP) is the top level directory containing VFP.exe and associated files;

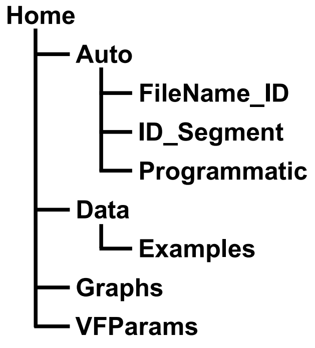


Figure 1. Directory structure created at program installation.

Sub-directory **Data** is the default suggested by the program for saving Data, SD and Regression files and **Examples** contains 29 example Data, SD and Regression files and one MS Excel file that are installed with the program. **Graphs** and **VFParams** are the defaults suggested for storing Graph files and exported variance function parameters respectively. **Auto** and its subdirectories contain example files illustrating how to implement automated data processing. You can add to or replace this subdirectory structure with whatever is necessary for your file organisation. The top level Home directory and *any* subdirectory structure beneath it is defined as the Program Data Area. The program incorporates browsers which are designed to assist quick identification of data and graphs but, in the interests of speed, the scanning range of the browsers is restricted to the Program Data Area. It is therefore strongly recommended that all data and graphs are saved to disk within this area.

User Preferences File

All user defined options and design settings are saved into the program user preferences file, VFP.usr, which is located in the program Home directory. If you wish to transfer program settings from one computer to another just copy VFP.usr to the program home directory on the target computer.

Auto-processing File

User defined automation Pointers, Protocols and other settings are saved into the program auto-processing file, Automate.usr, which is located in the program Home directory (refer to on-line help topic Contents > Automation > Automated Data Processing – Overview). In principle, it is possible to transfer auto-processing settings from one computer to another by just copying Automate.usr to the program home directory on the target computer. However, Pointers and Protocols rely heavily on directory settings (file input and output locations) and they will therefore fail (or at least require modification) if the directories do not exist on the target computer.

Graph Templates

The program allows the creation and maintenance of an unlimited number of graph templates. A template defines the colours, fonts, line widths, reference line style, tick mark lengths, aspect ratio, decimal separator and fixed text elements (axis labels, histogram, error and regression plot descriptors and Bland-Altman plot labels) that will be applied to a graph. Templates largely define the “artistic” appearance of a graph but they can also be used to switch fixed text elements from one language to another. All graphs dialogs contain a drop down list of currently available templates. Just select from the list to switch from one template to another depending on the output target (or language) for the current graph (printer, word processor, presentation graphics application, etc.) The idea is to avoid the need for constant fiddling with colours, fonts, etc. by pre-defining various combinations of settings in the form of templates. Several example templates are installed with the program (see on-line help topic Contents > Graphs > Graph Templates). They are samples only and can be renamed, modified or deleted.

User defined templates are stored in the program user preferences file VFP.usr (see above). However, one or more templates can be written to a separate disk file (.VFT extension), or read from a .VFT disk file. This can be used to provide separate backup for templates or to transfer selected templates from one computer to another without disturbing any other settings in the target user preferences file.

Decimal Separator

The program allows only a period or a comma. It defaults to a period if the system decimal separator is neither of these characters. The system decimal separator can be changed while the program is running; all summaries and graphs update immediately.

Screen Settings

The program assumes a screen resolution of at least 1280 x 768 pixels. Windows allows display sizing to be adjusted from the usual setting of 96 dpi (dots per inch) with an increase in dpi setting producing an increase in text size and *vice versa*. At a high screen resolution all program dialogs have been verified to display correctly up to a setting of 144 dpi (150% of normal) but at lower resolutions an increase in dpi setting can result in some dialogs not being completely visible (accessible) and/or parts of some labels not being visible. Note that the program provides Designers for all data summaries and these allow selection of different font sizes for screen display and for output. If necessary, select relatively large font sizes for screen display and smaller sizes for printing or clipboard output.

Scrolling through Summaries and Displays

Several dialogs contain summaries of data or files. In some cases the summary window can be focused meaning that one or more lines of the summary can be selected. In other cases the display is static and cannot be focused (eg. print preview displays). The consistent rule throughout the program is that whenever a summary or dialog has scroll bars it can be independently scrolled via the mouse wheel or

Alt+movement keys

where movement keys means Page Up, Page Down, Home, End or arrows. When a summary has focus it can be scrolled by movement keys alone (the highlight indicating the selected line moves), but the mouse wheel or Alt+movement keys can still be used to scroll it independently of the highlight position.

The Variance Function

The estimation method (5) used in this program takes full account of unequal replication in the data points which implies that the program can be used with any combinations of replicated data. The program was based initially around the "standard" 3-parameter variance function

$$\sigma^2 = (\beta_1 + \beta_2 U)^J$$

which is always monotone over the range of the mean values; either an increasing function of the mean over the entire range, or a decreasing function of the mean. Although very flexible, the 3-parameter function is capable of reducing to a straight line (ie. J = 1), including a horizontal straight line (ie. $\beta_2 = 0$). That does not mean that it will always do so. In cases of approximately constant variance it may also exhibit a sharp "hook" at one end or the other (see Fig. 2) which simply indicates that the model is over-parameterized (too flexible) in these situations. The program tests for the "hook" effect, and also for cases where the 3-parameter function has reduced to virtually a straight line, and in either case it substitutes the straight line variance function,

$$\sigma^2 = \beta_1 + \beta_2 U$$

as the default (eg. Fig. 3).

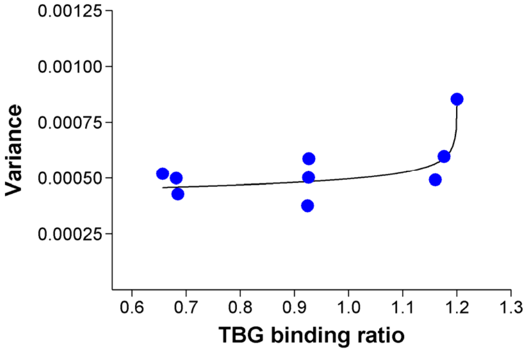


Figure 2. Example file TBGQC_93.VFD. TBG binding ratio is a T3-uptake test in which results have been converted to a ratio by dividing by the uptake for a normal serum pool. QC results from the 1993 calendar year exhibited a sharp "hook" at the upper end.

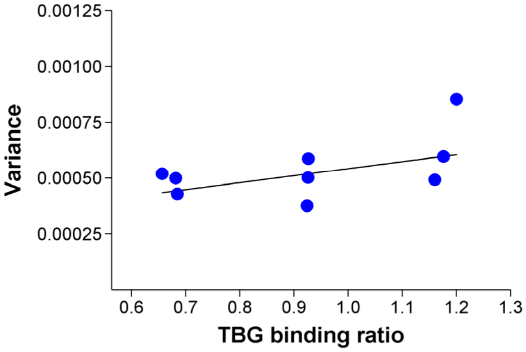


Figure 3. Example file TBGQC_93.VFD. Straight line variance function which is substituted when the "hook" effect is detected. Note that these data have markedly unequal replication. In particular, the top right data point was influential in producing the "hook" but has low replication (importance).

Figure 2 illustrates the failure of the highly flexible standard 3-parameter function when variance is approximately constant. However, at the other extreme, there are occasional data sets where the variance relationship shows a small but distinct U-shaped turning point near the assay detection limit and in these cases the standard function fails because of insufficient flexibility. Starting in program version 9.0, variance functions have been arbitrarily classified into two groups, "flexible" and "simple". The flexible functions (3 or 4 parameters) are intended for large change in variance situations, typified by immunoassay data, while the simple functions (1 or 2 parameters) are intended for general biochemistry tests or when experience suggests that a simple variance model might be indicated (eg. Fig. 2). One important distinction should be noted. When estimating a flexible function, the program can be configured to evaluate the available functions and present the most appropriate for the current data set (eg. automatically substituting a straight line as illustrated in Figs. 2, 3). In contrast, the user must select the most appropriate simple function.

Flexible Variance Functions

Figures 2, 3 illustrate how the program responds at one extreme. Figure 4 illustrates a rare U-shaped variance turning point near the assay detection limit. The 4-parameter function,

$$\sigma^2 = \beta_1 + \beta_2 U + \beta_3 U^J$$

was introduced in program version 8.0 to provide for this data pattern.

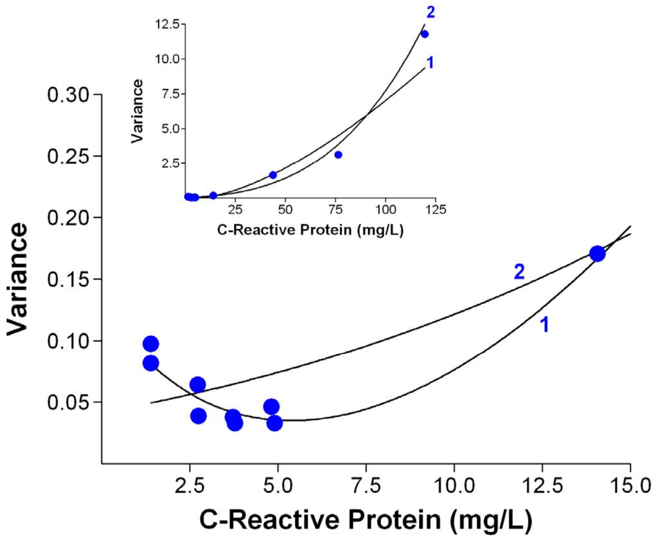


Figure 4. Example file CRPEval.VFD. These C-Reactive Protein data (*6*) consist of day-to-day replications of 12 evaluation specimens (265 degrees-of-freedom in total) and provide an excellent example of a variance turning point near zero. The data were kindly made available by Jill Tate, Royal Brisbane and Women's Hospital, Brisbane, Australia. The main part of the graph is a blow-up of the low concentration region while the inset shows the full range view. The 4-parameter and standard 3-parameter functions are labelled 1 and 2 respectively. The turning point 4-parameter function reduced the goodness-of-fit log likelihood ratio by nearly 80% relative to the monotone 3-parameter function.

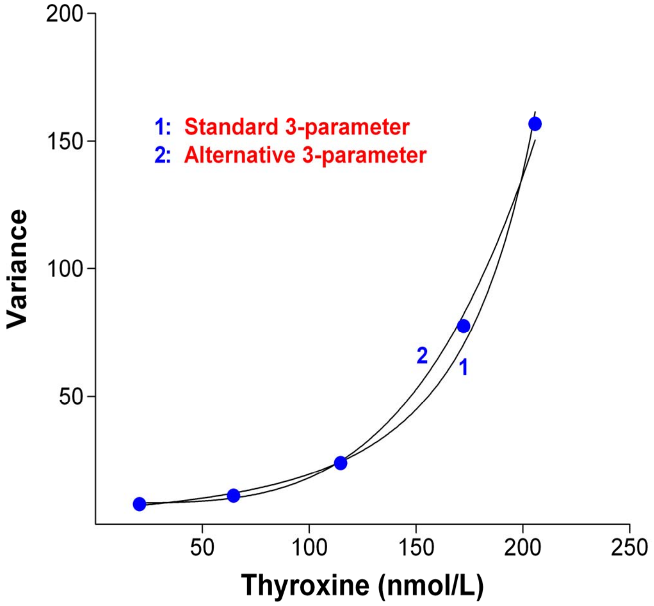


Figure 5. Example file T4QC.VFS.

Sampling experiments (*7*) using real immunoassay data showed that the alternative monotone 3-parameter variance function,

$$\sigma^2 = \beta_1 + \beta_2 U^J$$

has superior curvature properties at the low end of the range (relative to the standard 3-parameter function), and therefore offers potentially improved fits when monotone data are concentrated near the assay detection limit. This function has previously been suggested as a model for the immunoassay response-error relationship (*8*). Figure 5 shows a comparison of the standard and alternative 3-parameter functions and illustrates the typical patterns observed in the sampling experiments, ie. the standard function has superior curvature properties in the middle and upper parts of the range while the alternative function exhibits superior curvature at the low end (see also Fig. 14 on page 12). These monotone 3-parameter functions should be regarded as complementary rather than competitors.

Figures 2 – 4 show that no single function in the flexible group provides for all situations. Starting in program version 8.0, the aim is to identify and present the most suitable function for any particular data set. Users can be confident that program default settings will produce sensible, objective decisions. However, the program also allows considerable control over the criteria used in the selection process and this includes suppressing estimation of either or both of the 4-parameter and alternative 3-parameter functions. Access the dialog for defining preferences via a toolbar speed button or the main menu item Calculations & Plots > Variance Function Estimation Options > Flexible Function. A description of the issues can be found by clicking the Help button on the preferences dialog itself or by directly accessing on-line help topic Contents > Variance Function Options > Flexible Variance Function Options. Following variance function estimation a full description of the evaluation process is included in an estimation summary which can be viewed by clicking a speed button on the program toolbar or via the main menu item Calculations & Plots > Estimation Summary. It is strongly recommended that you accept the function suggested by the program (eg. the straight line and 4-parameter functions in Figs. 3, 4 respectively), but in virtually all cases you are given the option to select the standard 3-parameter function in graphical output, eg. for comparison purposes. It is worth mentioning that while the program default setting is to estimate the 4-parameter function, this requires considerable computing time when large numbers of data points are involved (eg. thousands of duplicates, as illustrated on following pages). Unfortunately, the potential gain is very small because of the rarity of turning points. Suppressing estimation of the 4-parameter function saves time when experience shows no evidence of turning points or when data are not concentrated in the vicinity of the assay detection limit. It can be switched on again at any time if and when indicated.

Simple Variance Functions

Five simple 1 or 2-parameter variance models were incorporated into the program in version 9.0;

$\sigma^2 = \beta_1$ (constant variance)

$\sigma^2 = \beta_1 U^2$ (constant CV)

$\sigma^2 = \beta_1 + \beta_2 U^2$ (mixed constant, proportional variance model)

$\sigma^2 = (\beta_1 + \beta_2 U)^K$ (constrained power model – K fixed at a user defined value)

$\sigma^2 = \beta_1 + \beta_2 U^K$ (alternative constrained power model – K fixed at a user defined value)

The constant variance model does not strictly qualify as a variance function because it implies no functional relationship between variance and the mean. However, constant variance over a range of mean values implies a non-constant relationship between CV and the mean which can be relevant in a medical laboratory context. Constant variance or constant CV relationships are frequently assumed for medical laboratory data, eg. the Bland and Altman method (*9*) for assessing the concordance of paired results from two analytical tests. Rocke and Lorenzato (*10*) proposed the mixed constant, proportional model. The two power models are equivalent to the flexible 3-parameter functions except that parameter J is replaced by a fixed exponent, K, whose value is defined by the user. Setting K = 1, for example, reduces both models to a straight line. As previously suggested (*7*), theoretical considerations, or analysis of large quantities of data, might produce a characteristic value of K for a particular test (particularly in an immunoassay context) and that value could be used thereafter to reduce sampling variation.

Figure 6 illustrates a set of arterial glucose intra-day (repeatability) precision data. The flexible standard 3-parameter function (plotted in terms of SD versus mean in this case) predicted a 2.69-fold increase in SD over the data range (7.25-fold increase in variance). However, simple inspection of the data does not support a trend and certainly not a systematic 2.69-fold increase in SD. Imposing a constant variance model seems appropriate. In essence, Fig. 6 is a less spectacular example (cf. Fig. 2) of the failure of a highly flexible variance function when change in variance is small or non-existent.

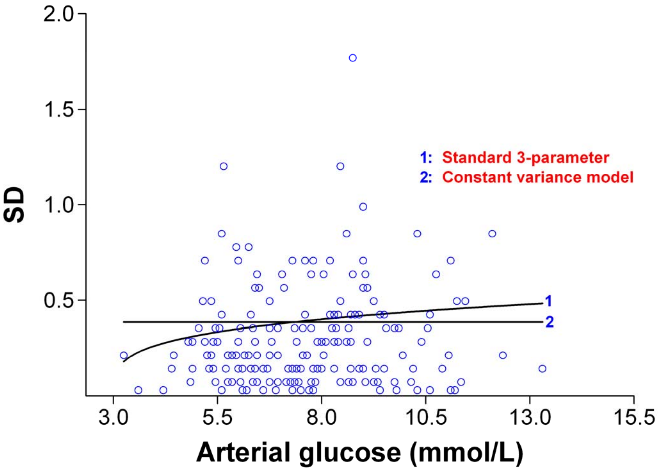


Figure 6. Example file Glucose.VFD. The data are 206 duplicate measurements of arterial glucose (206 degrees-of-freedom) where samples were drawn from patients in an intensive care setting. Measurements were made using the Precision PCx instrument (Abbott Laboratories, Abbott Park, Illinois, USA). The data were kindly provided by Peter Watkinson, The John Radcliffe Hospital, Oxford, UK. The flexible 3-parameter function reduced the goodness-of-fit log likelihood ratio by only 2% relative to a constant variance model.

The new Multi-estimation option introduced in program version 17.0 allows painless evaluation of the simple models prior to selecting the most appropriate. The constant variance and constant CV models are likely to be most useful when the data are well removed from zero concentration (eg. Fig. 6), while the mixed model of Rocke and Lorenzato (*10*) and the constrained power models are likely to be preferable when the data extend down to low concentrations.

Extrapolation

In common with any empirical method, the aim is to summarise and interpret the data that are actually available. Caution must be exercised when extrapolating outside the range of the data. An obvious example is estimating the SD at zero (SD[0]), which is often of special interest in an immunoassay context. If a reasonable number of data points or replications are located within a *few* SDs of zero it is perfectly acceptable to make the small extrapolation of the variance function necessary to predict SD[0] (set U = 0 in the fitted variance function equation then take the square root). This type of estimate has the advantage that it can summarise real operating conditions and data from real (clinical) specimens, as opposed to short-term replication experiments on what is sometimes, by necessity, an artificial zero specimen. However, a predicted SD[0] cannot be regarded as reliable, and may be grossly misleading, if the lowest mean values are many SDs from zero (this point is not always appreciated). Collection of appropriate data is the key. If appropriate data near zero are not readily available there is an immediate question mark over the *relevance* of SD[0]. See on-line help topic Contents > Overview > Limit of Blank and Limit of Detection for more information.

Variance Function Output Examples

Following pages show additional examples of program graphical output. All graphs in this PDF file were copy/pasted from the Variance Function Program to MS Word as bitmaps then scaled down inside Word. They illustrate some of the sample templates that are installed with the program. Any template could have been applied to any graph by just selecting it from a drop down list immediately prior to transferring the graph to Word. The graphs were constructed solely from the example Data, SD and Regression files that are installed with the program and could be easily recreated.

The program produces "standard" sized graphs for MS Word and MS PowerPoint. In the author's judgement, resolution at these "standard" sizes is acceptable for most purposes. If higher resolution is required the graph output dialog can be used to override the standard sizes with alternative user defined dimensions. The graph in Fig. 9 was created at twice the standard size; the rest were all standard size. The program default settings for exporting to Word specify no margin around the graph. However, when the graph had a coloured background this was overridden in the graph output dialog by specifying a 3% margin. The black borders around these graphs were added inside Word via the Borders and Shading option.

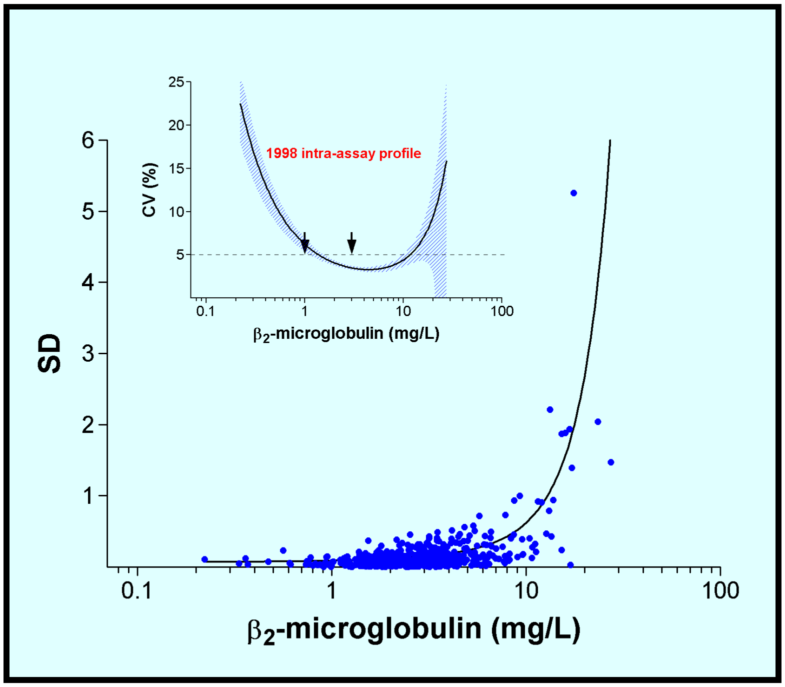


Figure 7. Example file B2mIntra_98.VFD.

Figure 7 illustrates the intra-assay variance function for a low throughput B2m RIA, estimated from clinical specimen duplicates for the 1998 calendar year (example file B2mIntra_98.VFD). The Data Type for this file is the rather strange looking **@Gb@S2-microglobulin (mg/L)** and is an example of the incorporation of Greek and subscript characters into a Data Type definition. As illustrated, the Data Type is translated appropriately in all graphical contexts. The inset shows the corresponding intra-assay imprecision profile which has been adjusted down to reflect duplicate measurement. The shaded area represents approximate 95% confidence intervals and the arrows define the assay reference range.

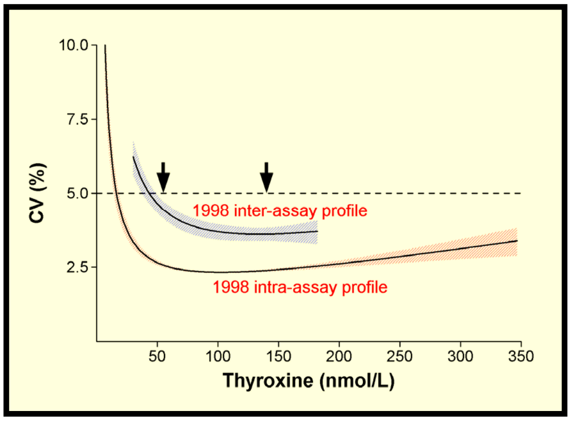


Figure 8. T4 RIA intra and inter-assay imprecision profiles for the 1998 calendar year, estimated from clinical specimen duplicates and QC specimen results respectively (example files T4Intra_98.VFD and T4QC_98.VFD). The intra-assay profile has been adjusted down to reflect duplicate measurement. The inter-assay profile intrinsically reflects duplicate measurement because the day-to-day QC results used to estimate the variance function were means of duplicates.

Figure 9 illustrates T4 RIA intra-assay profiles for 1998 by operator (example files T4S1_98.VFD through T4S8_98.VFD). All profiles have been adjusted down to reflect duplicate measurement. When confidence interval regions overlap, the program uses offset diagonal lines to create the effect of increasing density in direct relation to the number of overlaps at any particular point. When many overlaps occur, as in this example, higher resolution is required to achieve good quality. The program output dialog was used in this specific case to produce a graph that was twice the "standard" size for MS Word (ie. twice the resolution).

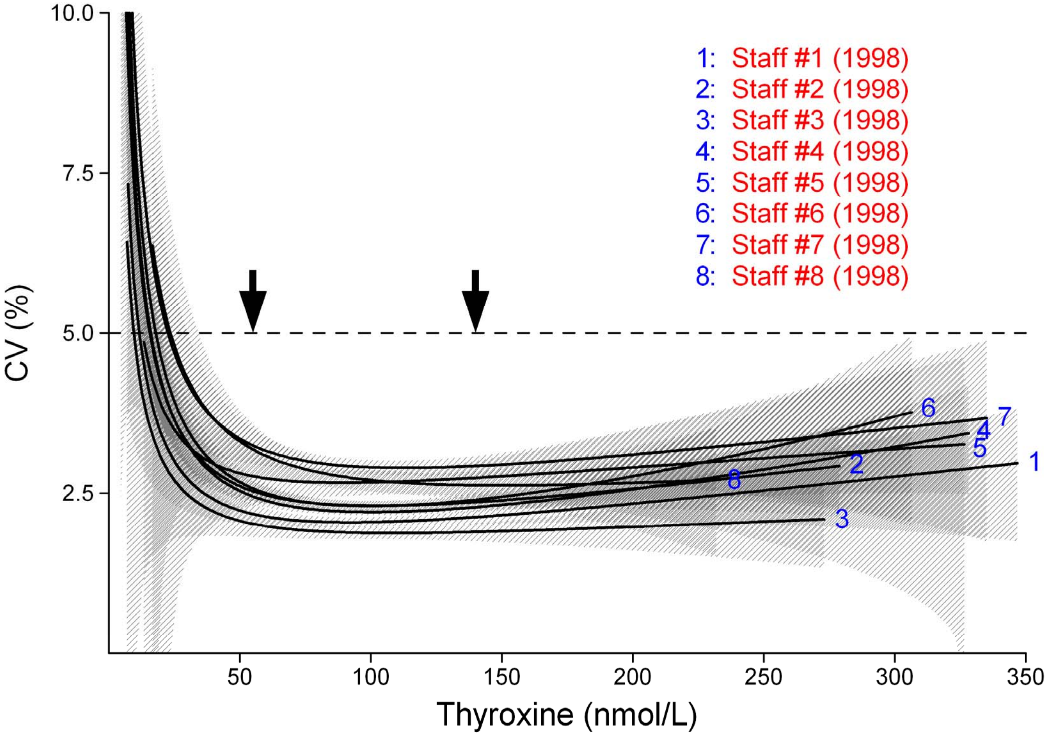


Figure 9. Example files T4S1_98.VFD through T4S8_98.VFD.

Figure 10 is a subset of Fig. 9. The order in which the data are plotted is determined by the order in which the variance functions are assigned to the graph. In this case the red data points produced by Staff #3 (assigned second) plot on top of the blue data. Data can be also be specified to plot over or under the variance functions (under in this case which allows the variance functions to be fully visualised).

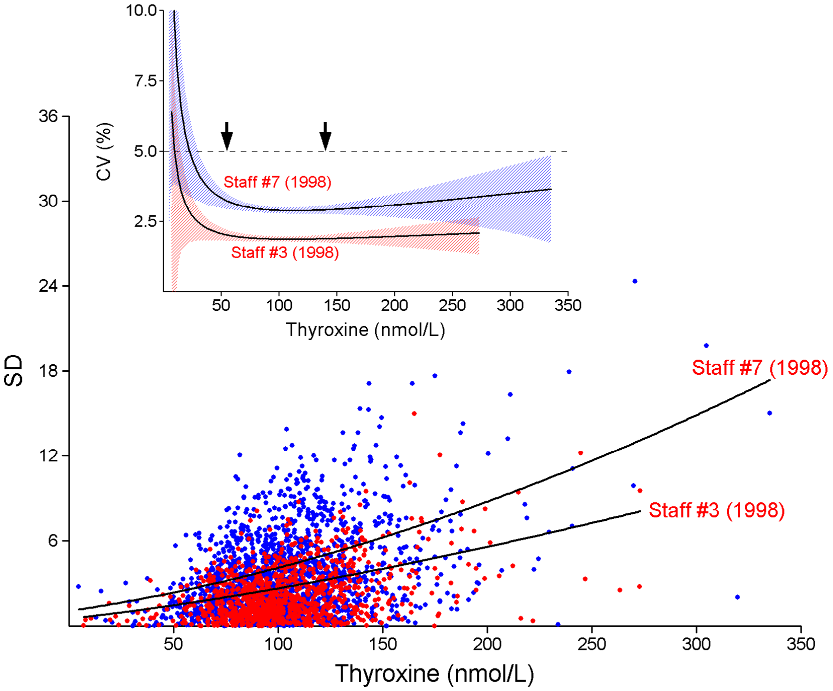
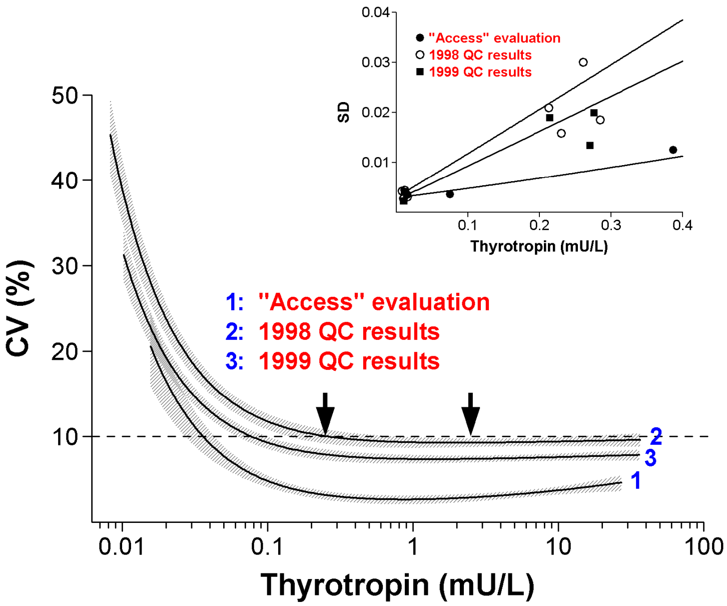


Figure 10. Example files T4S3_98.VFD and T4S7_98.VFD.



Systematic differences of the type illustrated in Figs. 9, 10 are not necessarily a property of manual assays only. In our experience (Fig. 11), comparable or larger differences can occur with automated instruments depending upon the conditions of use. It is not clear whether the TSH assay variability we have observed is solely a property of our “Access” instrument (Serono/Beckman/Coulter, Fullerton, CA, USA), or is a property of the “Access” instrument in general, or of automated instruments in general. The inset is a blow-up of the low concentration region. It illustrates very consistent (and excellent) reproducibility near zero, rapidly diverging at higher concentrations.

Figure 11. Example files TSHEval.VFD, TSHQC_98.VFD and TSHQC_99.VFD.

Variance Stabilisation

Moderate to wide range immunoassay data invariably violate the uniform variance assumption inherent in many parametric statistical analyses. However, Bartlett (*11*) showed that if the form of the relationship between variance and the mean is known, say $\sigma^2 = \phi(U)$, then transformation of the raw observations by,

$$T(U) \ = \ \int[\phi(U)]^{-½}\, dU$$

should stabilise variances to a value 1.0 (as a first order approximation). Assuming the "standard" 3-parameter variance function is a good general variance model for immunoassay data, then (as we noted in Ref. *1*) variance should therefore be stabilised by applying the following transformation to the raw replicate values (x),

$$T = [(1 - J/2)\beta_2]^{-1}[\beta_1 + \beta_2 x]^{1-J/2} \qquad (J \neq 2), \qquad \text{or} \qquad T = \log[\beta_1 + \beta_2 x]/\beta_2 \qquad (J = 2).$$

This transformation is available for any values of parameters β_1, β_2 and J. When a straight line function has been substituted (ie. J = 1), the transformation reduces to,

$$T = (2[\beta_1 + \beta_2 x]^{1/2})/\beta_2$$

Note that because the standard 3-parameter and straight line functions retain monotonicity across zero they also retain their transformation properties in the presence of zero and negative mean values. In contrast, the 4-parameter and alternative 3-parameter functions yield a simple integral (hence transformation function) only when parameter J has a small integer value (a vanishingly small frequency in practice). For the same reason, not all of the simple variance functions yield a transformation function. It is therefore strongly recommended that a flexible variance function is estimated prior to a variance stabilising transformation because this guarantees the availability of the standard 3-parameter function and/or the straight line function. The program automates the transformation process to a button click. Transformed data are written into a new file with exactly the same characteristics as the original file and full details of the transformation are written into the notes of the newly created file.

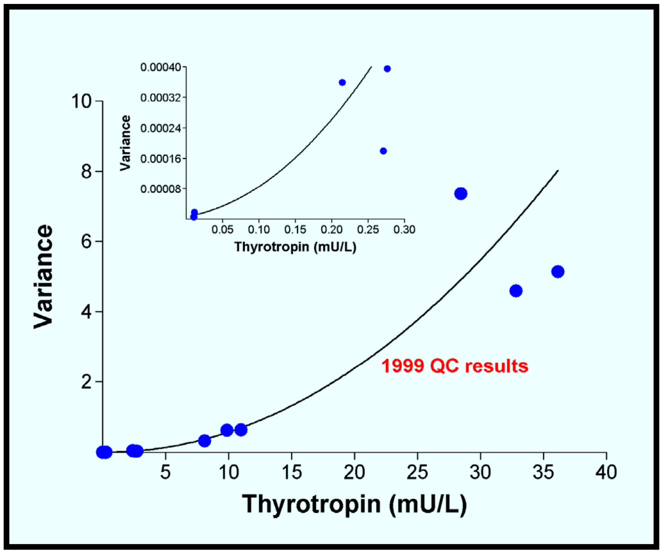


Figure 12. Example file TSHQC_99.VFD.

Figure 12 shows 1999 TSH QC data and fitted variance function The inset is a blow-up of the origin region. Although it may not be obvious visually, predicted relative change in variance for these data is > 750000-fold.

Figure 13 shows the data after transformation (example file TSHQC_99VF.VFD). The variance function fitted to these transformed data had a "hook" at the lower end and was replaced by the straight line variance function. Predicted relative change in variance was reduced from >750000-fold to 1.22-fold. The formulae given above are general variance stabilising transformations. We have found them to work extremely well in practice.

As it happens, the Fig. 12 relationship between variance and the mean was very close to a quadratic (variance function parameter J very close to 2.0), and this would cause a statistician to immediately think of a log transformation. Figure 14 (example file TSHQC_99Log.VFD) shows the result of the simple transformation,

$$T \ = \ \ln(x) + 10$$

where ln denotes natural logarithm and the arbitrary constant 10 was added to ensure all transformed results were positive. The simple log transformation worked extremely well for data at the four higher QC concentrations but it failed to pull the very low data into line.

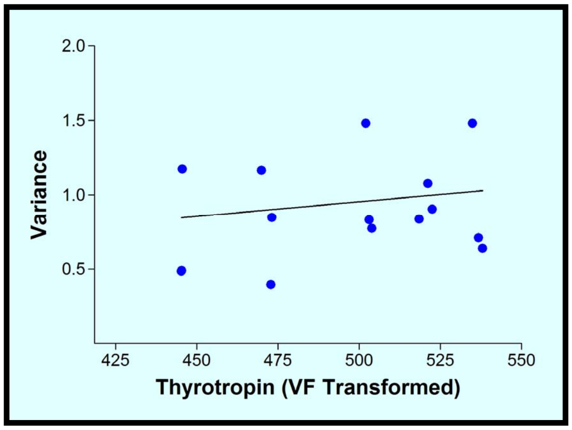


Figure 13. Example file TSHQC_99VF.VFD. The standard 3-parameter function exhibited a "hook" at the lower end the straight line variance function has been substituted.

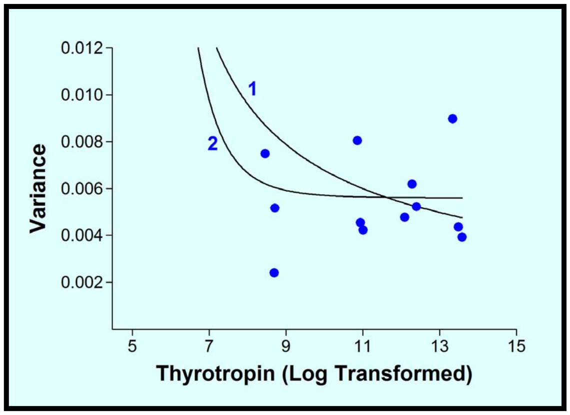


Figure 15. Example file TSHQC_99Log.VFD. Expanded view of the four higher QC concentrations from Fig. 14.

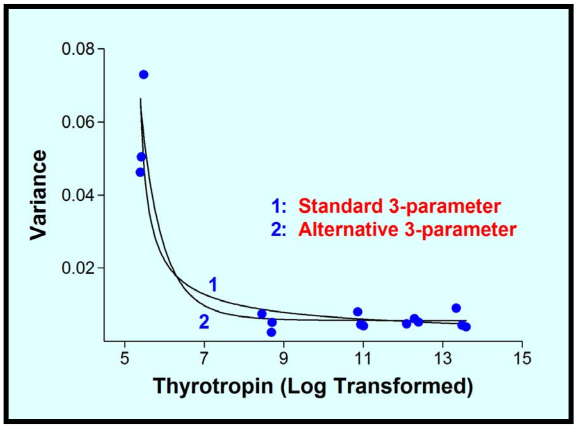


Figure 14. Example file TSHQC_99Log.VFD. The alternative 3-parameter variance function provided a better fit than the standard 3-parameter function.

Figure 15 is an expanded view of the four higher QC concentrations from Fig. 14 and illustrates the similarity, in this particular case, of results from variance function transformation (Fig. 13) and log transformation at concentrations well removed from zero. However, a log transformation assumes an underlying variance relationship,

$$\sigma^2 = \beta_1 U^2$$

which (incorrectly) predicts $\sigma^2 = 0$ at U = 0. Simple transformations like the logarithmic or square root will always fail when data extend down toward the detection limit of a measurement system. As a digression, the alternative 3-parameter function provided a better fit to the atypical data in Figs. 14, 15 because of its superior low-end curvature properties.

Histogram, Error and Biological Variation Plots

Levey-Jennings, Cusum and other control-oriented graphs are essential laboratory tools. Taken alone, however, they do not give much of an impression of the actual error of a measurement system. For example, two runs of QC results with CVs of 1% and 20% could have an identical appearance on Levey-Jennings plots. The program produces histogram plots of sets of replicates which represent measurement error as it would be perceived by consumers. It also produces error distribution plots where the error can be user defined (visual experimentation), real error predicted from variance functions, or a mixture of the two, at user defined mean values. The intention is to give an alternative visual perspective. Figures 16, 17 are examples of output. An unlimited number of plots of this type could be produced from any particular Data or SD file. The main points of interest are not the middle parts of the distributions (even low quality assays produce a majority of results relatively close to the underlying mean values) but the *tails* of the distributions where results have the potential to at very least cause confusion.

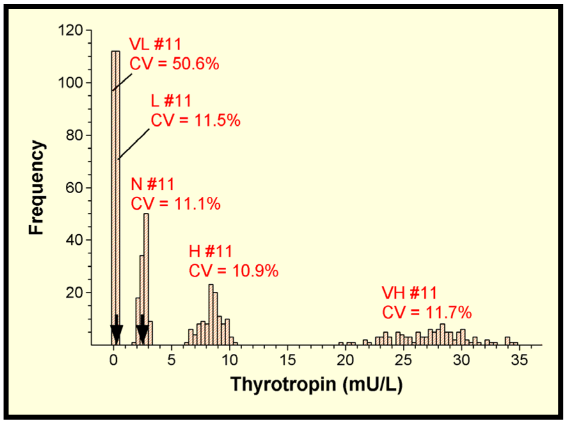


Figure 16. Example file TSHQC_98.VFD. QC lot #11 results (N=112), accumulated during the middle part of 1998.

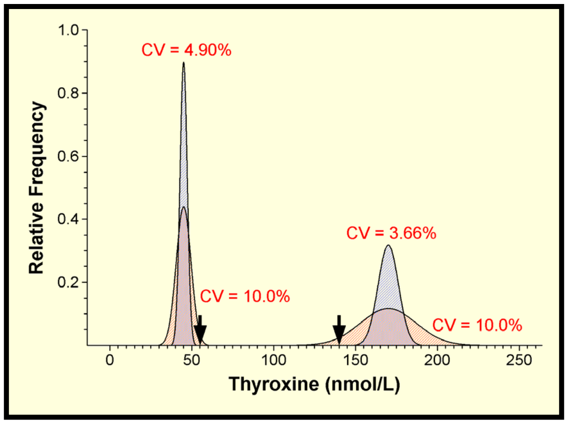


Figure 17. Example file T4QC_98.VFD. Inter-assay error predicted from the variance function versus hypothetical error (CV = 10%).

Histogram and error plots do not necessarily show the full picture. Hormones and other biological substances that are continuously secreted *invivo* have a natural day-to-day variation about a set-point (over and above cyclical rhythms). It is the variation consumers (clinicians) would see if sampling times and techniques were absolutely consistent and assays had zero error. Several groups have proposed biological variation as a basis for linking laboratory performance goals to clinical needs. For example, if the intra-individual biological variation of a hormone is B%, then a suggested laboratory goal is day-to-day assay CVs ≤(B/2)%. This limits additional variability due to assay errors to ~10%. C.G. Fraser and P.H. Petersen have published many papers on the subject. They also provide informative summary essays on the world-wide web (Refs. *12*, *13*). Ricos *et. al.* (*14*) provide a comprehensive database of biological variation values.

The program produces plots of intra-individual biological variation plus the *additive* effect of measurement error, where the error can be user defined (visual experimentation), real error predicted from variance functions, or a mixture of the two, at user defined mean values. Figure 18 demonstrates that the large discordance between evaluation performance of an automated TSH assay and subsequent routine performance (Fig. 11) does not make much difference for general clinical purposes because of the large biological variation of TSH (~20%). For this reason we did not reject the instrument. If we had been using the instrument to also measure T4, and assuming similar performance characteristics, a different outcome would have occurred (Fig. 19).

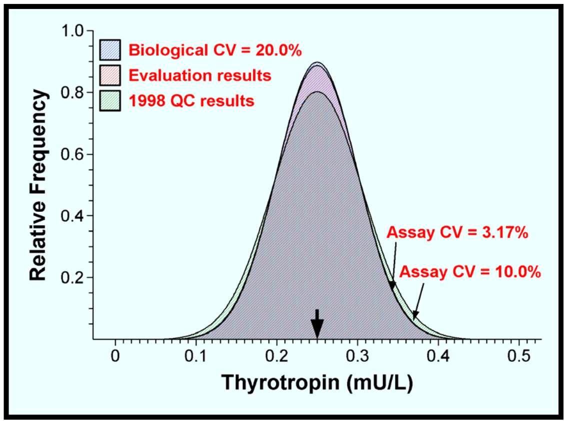


Figure 18. Files TSHEval.VFD and TSHQC_98.VFD.

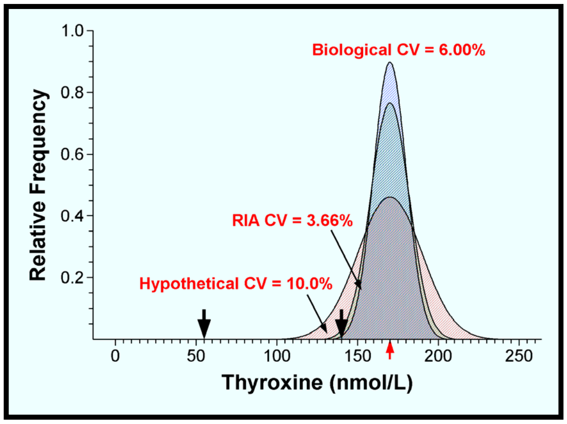


Figure 19. File T4QC_98.VFD and a hypothetical CV of 10%.

Regression Analysis

Figures 20, 21 illustrate graphical output from the program Regression dialog. Interested users should consult on-line help topic Contents > Overview > Regression. Ref. *15* provides numerical details including the advantages of significance testing using joint parameter confidence regions rather than slope, intercept confidence intervals.. Additional information is contained in the notes in example Regression files T4_Dups.VFR (illustrated below), TSH_Eval.VFR and T4_Dil.VFR.

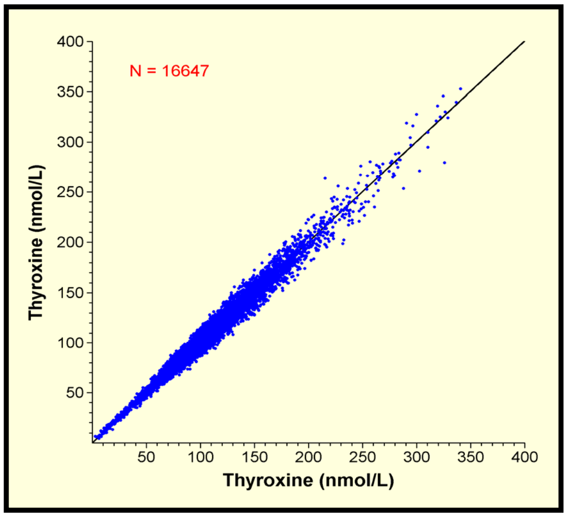


Figure 20. Idealised Deming (*16*) regression plot (file T4_Dups.VFR). The X, Y pairs are the duplicates originally in file T4Intra_98.VFD (see Fig. 8) which provided the X, Y weighting functions in this case.

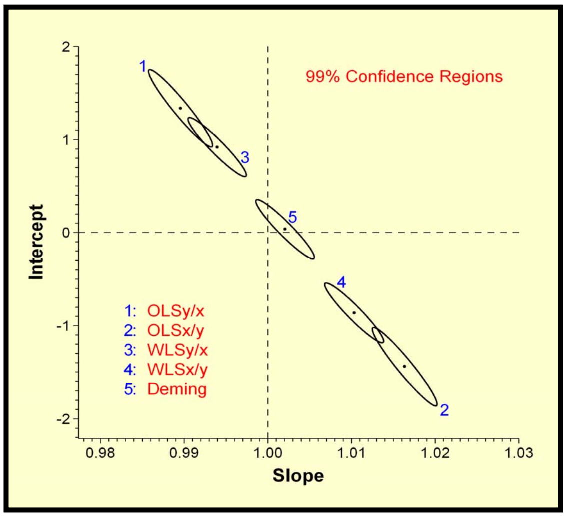


Figure 21. Joint parameter confidence regions from unweighted (OLS), weighted (WLS) and Deming (errors in X and Y) regressions on the data in Fig. 20.

Limit of Blank (LoB) and Limit of Detection (LoD)

The program incorporates four distinct ways of estimating LoB then uses a suitable variance function to estimate LoD and 95% confidence intervals as per ISO definitions (*17*). Interested users should consult on-line help topic Contents > Overview > Limit of Blank and Limit of Detection for the general principles involved and/or click the **Help** button on the LoB, LoD calculations dialog for information on practical issues including data and variance function requirements. Complete numerical details are given in Ref. *18*, but it should be noted that small biases reported in that paper were subsequently shown (*19*) to be artefacts thrown up by simulation experiments that lacked the necessary precision.

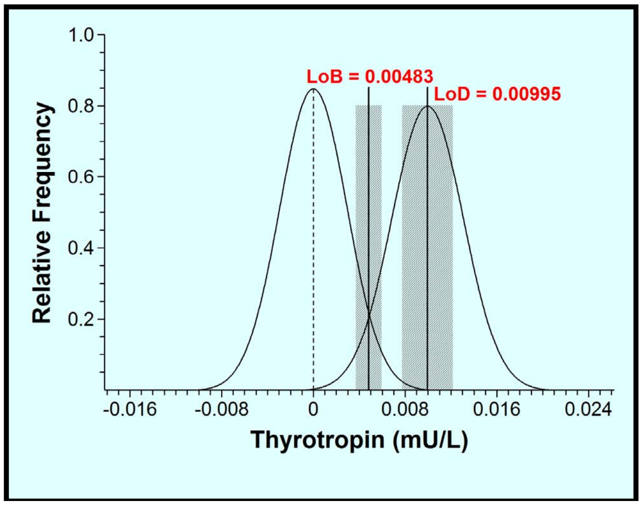
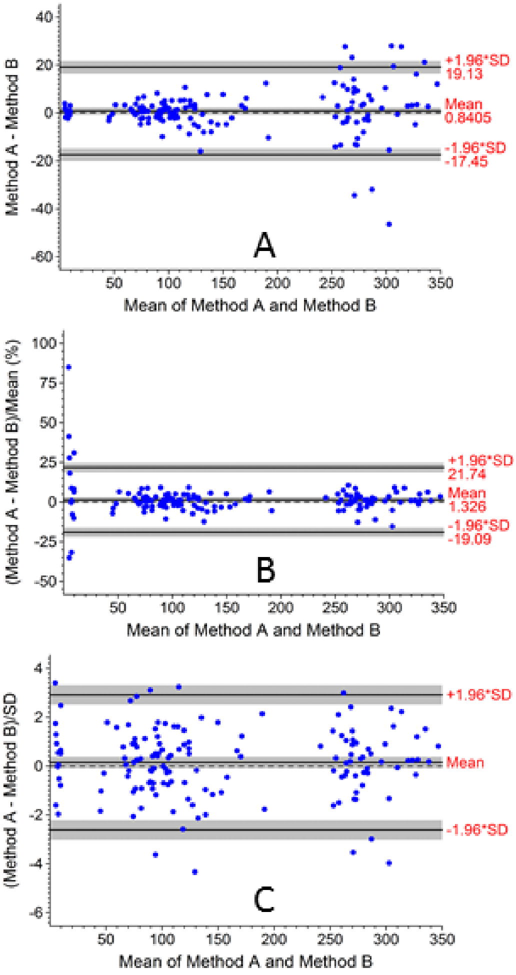


Figure 22. LoB, LoD results and graphical output estimated from example file TSHEval.VFD. The shaded areas are 95% confidence intervals.

At least four of the example files that are installed with the program (see Appendix D) are suitable for illustrating LoB, LoD calculations and graphical output but this does require the assumption that extrapolating each of the variance functions to zero yields a reliable estimate of the SD at zero and also that the distribution of blank specimen replicates is Gaussian. While the latter, in particular, requires independent verification the aim here is merely to illustrate the calculations in a painless way. Figure 22 illustrates graphical output from the LoB, LoD dialog calculated from the TSH TSHEval.VFD (see Fig. 11).

Alternative calculation examples are provided by the data contained in read-only MS Excel file, Blanks_Data.xls, which is installed into C:\VFP\Data\Examples along with the program example files. This Excel file also includes detailed instructions on how to use the data.

Bland-Altman Analysis



Hypothetical data to illustrate Bland-Altman analysis were extracted from example file T4Intra_98.VFD (see Figs. 8, 9 and 20) which contains 16647 T4 RIA duplicates. All duplicates with mean values < 10 nmol/L (N = 14) and with mean values > 250 nmol/L (N =50) were combined. The remaining 16583 duplicates were listed in the order they appear in the file and every 192nd duplicate was added to the 64 previously extracted (N = 150 in total). First and second replicate values were assigned as Method B and Method A respectively. Figure 23A shows the resulting Bland-Altman plot in terms of raw results and Fig. 23B shows the plot after adjusting differences by the factor 100/Mean (equivalent to Bland-Altman calculations after logarithmic transformation of individual pair values).

Traditional Bland-Altman analysis assumes either constant variance (Fig. 23A) or constant CV (Fig. 23B) and can therefore produce limits of agreement that are misleading or even meaningless when the data do not conform to either of those simplifying assumptions. Figures 23A and B are a mild example of that. More extreme examples occur whenever clinically important results are located near assay detection limits where CV rises sharply (and results from numerous immunoassays fall into that category).

Figure 23. Bland-Altman plots derived from 150 paired T4 RIA results (nmol/L). Raw pair differences (Panel A), differences adjusted by the factor 100/Mean (Panel B) and differences divided by predicted SD evaluated at corresponding mean values (Panel C). Shaded areas are 95% confidence intervals.

Hawkins (*20*, *21*) suggested estimating the Rocke and Lorenzato function (*10*) from the pair differences then normalising the pair differences by variance stabilising the A, B pair values. This program uses the alternative approach of dividing pair differences by predicted SD evaluated at corresponding mean values. In the case of these T4 results the standard 3-parameter function was used, as per (*22*), and the result is shown in Fig. 23C. Hawkins also adapted the Breusch-Pagan test (*23, 24*) as an objective measure of the homogeneity of the raw or transformed differences but it is hardly required for this example. Panel C differences are clearly homogeneous across the entire range of mean values and are a marked (if not spectacular) improvement on Panels A and B.

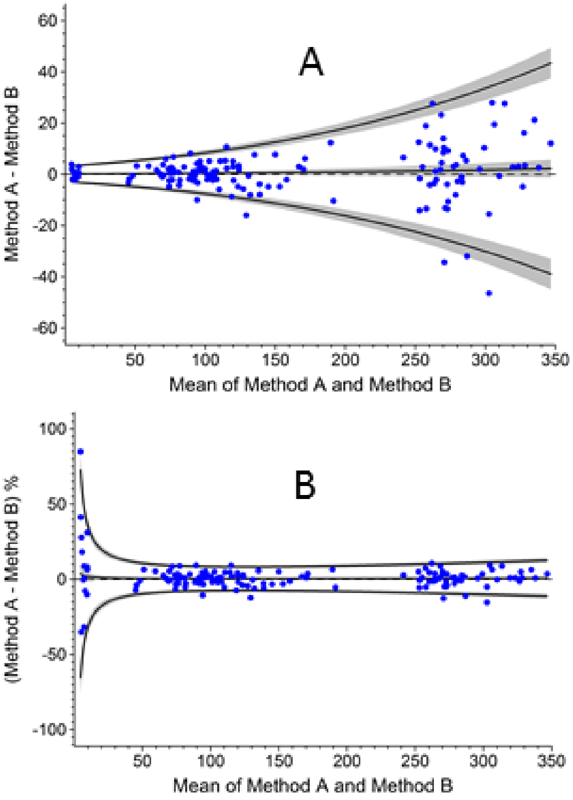


Figure 24. Data and results from Figure 23C back-transformed to the original measurement scale (Panel A) and expressed as percentages of mean values (Panel B). Measurement units are nmol/L.

Unfortunately, whereas the mean bias and limits of agreement have simple interpretations in Figs. 23A (raw measurement units) and 23B (percentages), the corresponding values in Fig. 23C have no obvious meaning. However, it is a simple matter to back-transform the Fig. 23C data and summary lines to either the original measurement scale or as percentages (illustrated in Figs. 24A and 24B, respectively). The *data* in Figs. 24A and B are identical to the data in corresponding panels of Fig. 23, but the back-transformed limits of agreement in Fig. 24 can be considered legitimate across the entire range of mean values, in contrast to those in Figs. 23A and B. In general, limits of agreement *are* concentration dependent, as illustrated in Fig. 24, unless the data just happen to conform to either of the constant variance or constant CV assumptions that traditional Bland-Altman analysis relies on.

The Bland-Altman module (introduced in program version 15.0) incorporates traditional Bland-Altman analysis (eg. Figs. 23A and B) but also includes data normalisation by any of the variance functions offered by the program (estimation and evaluation was automated to a button click in program version 16.0). Back-transformation (eg. Fig 24) is accompanied by automatic presentation of bias and limits of agreement values (expressed in raw units and as percentages), plus the option of blow up views, at user defined mean values. See on-line help topic Contents > Overview > Bland-Altman Analysis for more information.

Appendices

A. Program Delivery

The Australasian Association of Clinical Biochemists has kindly allowed the program installation file to be downloaded from its website (www.aacb.asn.au). Proceed to Resources > Tools and Links which contains a link to the Variance Function Program. Click the installation file link to download VFP190Install.zip (10897 Kb) and extract the installation file VFP190Install.exe (11040 Kb). Double click it to launch the installation process.

B. Program Installation

Initial Installation

Install version 19.0 by double clicking file VFP190Install.exe and following the prompts. VFP.exe (12385 Kb), on-line help file VFPHelp.chm (852 Kb), user preferences file VFP.usr (13 Kb) and automation file Automate.usr (6 Kb) are installed into directory C:\VFP and a shortcut icon is placed on the desktop.

Upgrading from a previous version

It is strongly recommended that you uninstall any previous version. The uninstall will remove any existing example files (which will be replaced) but any user generated data and graph files will not be effected. However, in order to retain your current program settings, graph templates and automation settings, it is important that you copy the current user preferences file (VFP.usr) and automation file (Automate.usr) to a temporary location (eg. C:\Temp) before commencing the uninstall/install process. When version 19.0 is installed copy the saved VFP.usr and Automate.usr to C:\VFP, the version 19.0 home directory (see **User Preferences File**, **Auto-processing File** and **Graph Templates** on pages 3, 4). If the previous version was located in C:\VFP then you will have immediate access to the data and graph files created by the previous version. If not, data and graph files created with the previous version should be copied to subdirectories under C:\VFP in order to make them accessible to the version 19.0 data and graph browsers (see **Program Data Area** on page 3).

If you choose to install version 19.0 alongside a previous version then, if necessary, you must rename the existing home directory to something other than C:\VFP because version 19.0 will be automatically installed into C:\VFP and will overwrite any existing programs. After installing version 19.0 you should copy the previous VFP.usr and Automate.usr files to C:\VFP if you wish to retain the program settings, graph templates and automation settings defined with the previous version. Likewise, data and graph files accumulated with the previous version should be copied to subdirectories under C:\VFP in order to make them accessible to the version 19.0 data and graph browsers (see **Program Data Area** on page 3).

C. Program Removal

Use the Add/Remove Programs icon (or equivalent) in the system Control Panel to remove the program icon and all installed files. Any files created after program installation will not be included in the clean up operation and must be deleted manually (eg. user generated Data, SD, Regression and Graph files).

D. Example Data, SD and Regression Files

The following 29 example Data (.VFD), SD (.VFS) and Regression (.VFR) files are installed with the program into directory C:\VFP\Data\Examples. The text to the right is the descriptive header contained in each file.

T4Intra_98.VFD	T4 RIA: Clinical specimen duplicates for 1998 calendar year
T4S1_98.VFD	T4 RIA: Staff #1 clinical specimen duplicates for 1998 calendar year
T4S2_98.VFD	T4 RIA: Staff #2 clinical specimen duplicates for 1998 calendar year
T4S3_98.VFD	T4 RIA: Staff #3 clinical specimen duplicates for 1998 calendar year
T4S4_98.VFD	T4 RIA: Staff #4 clinical specimen duplicates for 1998 calendar year
T4S5_98.VFD	T4 RIA: Staff #5 clinical specimen duplicates for 1998 calendar year
T4S6_98.VFD	T4 RIA: Staff #6 clinical specimen duplicates for 1998 calendar year
T4S7_98.VFD	T4 RIA: Staff #7 clinical specimen duplicates for 1998 calendar year
T4S8_98.VFD	T4 RIA: Staff #8 clinical specimen duplicates for 1998 calendar year
T4QC_98.VFD	T4 RIA: QC specimen results for 1998 calendar year
T4Intra_99.VFD	T4 RIA: Clinical specimen duplicates for 1999 calendar year
T4S1_99.VFD	T4 RIA: Staff #1 clinical specimen duplicates for 1999 calendar year
T4S4_99.VFD	T4 RIA: Staff #4 clinical specimen duplicates for 1999 calendar year
T4S9_99.VFD	T4 RIA: Staff #9 clinical specimen duplicates for 1999 calendar year
T4QC_99.VFD	T4 RIA: QC specimen results for 1999 calendar year

TSHEval.VFD	"Access" TSH ILMA: Evaluation results (Mar – May 1997)
TSHQC_98.VFD	"Access" TSH ILMA: QC specimen results for 1998 calendar year
TSHQC_99.VFD	"Access" TSH ILMA: QC specimen results for 1999 calendar year
TSHQC_99Log.VFD	"Access" TSH ILMA: Log transformed QC specimen results for 1999 calendar year
TSHQC_99VF.VFD	"Access" TSH ILMA: VF transformed QC specimen results for 1999 calendar year

B2mIntra_98.VFD	B2m RIA: Clinical specimen duplicates for 1998 calendar year
TBGQC_93.VFD	TBG binding ratio: QC specimen results for 1993 calendar year
CRPEval.VFD	"IMMAGE" CRP ILMA: Evaluation results that generated a variance turning point
T4QC.VFS	T4 RIA: Collection of QC results used to compare 3-parameter variance functions
Glucose.VFD	Precision PCx: Intra-day precision of arterial glucose measurement
T4_Dups.VFR	An idealised regression: T4 RIA duplicates from example file T4Intra_98.VFD
TSH_Eval.VFR	TSH method comparison: Serono IRMA (X) vs Sanofi "Access" ILMA (Y)
TSHIRMA.VFD	Serono TSH IRMA: QC specimen results March through May 1997
T4_Dil.VFR	T4 RIA: Serial doubling dilutions

These files can be freely modified by experimentation. If you wish to retain access to the originals it is suggested that you copy them to a separate archive directory (outside the Program Data Area). In addition, a read-only MS Excel file, Blanks_Data.xls, is also installed into C:\VFP\Data\Examples. It contains blank specimen data and instructions for illustrating four different LoB estimation methods in the program Limit of Blank and Limit of Detection dialog.

E. Auto-processing Example Files

Auto-processing file Automate.usr is installed into C:\VFP. It contains three example Pointers, four example Protocols and default auto-processing event log settings. Six graph (.VFG) files, a read-only Word document (AutoQC.doc) and a graph template file (IntQC.VFT) are installed into directory C:\VFP\Auto.

Directory C:\VFP\Auto\FileName_ID will contain 22 input data files (text files) with names of the type Pr_T4QC_Inter^……..Txt. Directories C:\VFP\Auto\ID_Segment and C:\VFP\Auto\Programmatic will also contain 22 input data files (text files) with names InputA_nn.Txt and InputB_nn.Txt, respectively (nn = 01 – 22).

In brief, the three groups of 22 input files contain identical data but illustrate different ways of communicating processing instructions to the program. An auto-processing run on any of the three groups will generate the same set of 22 Data (.VFD) files in directory C:\VFP\Auto. The six graph (.VFG) files mentioned above have been configured to display the variance functions contained in the 22 .VFD files. AutoQC.doc illustrates how the graph output is organised in this laboratory and IntQC.VFT contains the graph template used.

Consult on-line help topic Contents > Automation > Automated Data Processing – Overview for a detailed explanation.

F. History

- Version 1.0 (September 1990; Ref. *25*) was written in Fortran 77 for the MS-DOS environment. It focused on the estimation and plotting of immunoassay imprecision profiles.
- Version 2.0 (September 1999) was the initial Windows version, written in Borland's Object Pascal (not distributed).
- Version 3.0 (July 2000; Ref. *26*) was an improved Windows version.
- Version 4.0 (May 2001; Ref. *27*) added automated variance stabilising transformation, error plots and biological variation plots.
- Version 5.0 (May 2005) added a CV Calculator, automatic data splitting and enhanced options for selectively switching data and variance functions on/off in graph views. This upgrade also incorporated the opening of program data and graph files in Windows Explorer by double clicking or drag-and-drop.
- Version 6.0 (November 2005) added a parameter export dialog which facilitates variance function graphical work or other analyses in external applications, but the main change was the incorporation of completely automated data processing, intended for moderate to large scale repetitive analyses such as internal QC or simulations.
- Version 7.0 (June 2006) added a regression spreadsheet (paired X, Y values) and a regression dialog which calculates unweighted, weighted and Deming (errors in X and Y) regressions.
- Version 8.0 (July 2007) incorporated two additional variance functions; a 4-parameter function that provides for rare variance turning points near zero and an alternative 3-parameter function that has greater low end flexibility. Revision 8.2 made some important changes to regression calculations.

- Version 9.0 (July 2009) incorporated a group of five simpler 1 or 2-parameter variance functions, such as a constant variance model and a constant CV model, aimed at laboratory tests other than immunoassays.
- Version 10.0 (April 2011) upgraded the program to be Windows 7 compatible. Now requires Windows 2000 or later, but remains fully backward compatible. On-line help was translated from WinHelp to HTML help.
- Version 11.0 (January 2013) increased data capacity by a factor of five, significantly increased the number of available Undo/Redo steps during spreadsheet data editing and updated variance and SD plots to allow a logarithmic scale on the Y-axis.
- Version 12.0 (December 2014) introduced a new Limit of Blank and Limit of Detection dialog for estimating detectability characteristics according to ISO guidelines (*17*), and incorporated program-wide touch screen gestures to facilitate several of the most common graphical manipulations.
- Version 2016 (August 2016) automated the processing of blank specimen data by allowing blanks data to be either incorporated into variance function estimation, or imported as a separate body of data that can be saved with the file.
- Version 14.0 (June 2017) incorporated Bartlett's test for homogeneity of variance, plots of regression weighting functions, sorted lists of blank specimen replicates and a number of cosmetic changes to the user interface.
- Version 15.0 (July 2018) incorporated Bland-Altman analysis including pair difference normalisation by variance function. Input/output of replicated data was extended to include Excel style Cases and Values format. Convergence criteria for estimated variance function parameters were tightened.
- Version 16.0 (May 2019) updated aspects of Bland-Altman analysis including, in particular, automating the estimation and evaluation of normalising variance functions.
- Version 17.0 (May 2020) added a multi-estimation speed button which allows estimation of all variance functions with a single button click, increased the allowed range of user defined exponent values and updated paired X, Y data requirements to allow zero and negative values.
- Version 18.0 (November 2021) relaxed precision data requirements by allowing zero and negative replicate values and zero and negative set mean values, subject to the requirement that values must be plausible observations from the detection limit region of a measurement system.

G. Changes: Version 18.0 to Version 19.0

1. A minor program update. After updating the program development environment (Embarcadero's RAD Studio) to the latest Windows 11 version the program was migrated and rebuilt to ensure that it performs correctly under Windows 11. Intended to provide a bit of future-proofing. Version 19.0 was also verified to successfully install and run back to at least Windows 8.2.
2. The program offers exporting of numerical details associated with variance function estimation. Almost all program analyses and graphs are based solely on these summary details (exceptions are variance stabilising transformations and histogram plots which also require access to raw replicate values). The aim is to allow users to perform additional analyses or graphical work in external applications and that includes using parameter exporting in conjunction with the program Automation module to perform simulations (this has been used to perform extensive statistical experiments as part of program testing, eg. Refs. *18, 19*). Output was previously restricted to either a text file or to the Windows clipboard with numerical values in scientific format. Data for export can now be examined in a viewing window and users are also given the option of either scientific or decimal format. Decimal format is probably preferable for viewing purposes but scientific format is strongly recommended when calculations are intended.

Consult on-line help topic Contents > Appendix for details of updates across all program versions.

H. Version 18.0 Bug Fixes

1. The version 18.0 upgrade to allow zero and negative replicate values required substantial changes to both the program and to on-line help. A few paragraphs in on-line help were not modified appropriately. Now updated.

2. The Windows Font selection dialog provided as part of Embarcadero RAD Studio reliably implements the selected font, style and size chosen by the user but, unfortunately, if the selected style is not one of the standard styles (regular, **bold**, *italic* (aka *oblique*) or ***bold italic*** (aka ***bold oblique***)) the dialog returns a modified font name that cannot be interpreted by the program. The font is successfully applied but if the Font dialog is recalled in these cases the font name and style are missing from the dialog edit boxes. This is very irritating and the work-around is to limit font styles to the four standard styles mentioned above, ie. selected styles are

converted to the nearest standard equivalent, eg. "Ultra light" is converted to regular, "Extra Bold" is converted to bold, etc. Not a particularly serious restriction but necessary to ensure that font details are retained for future inspection.

3. Estimation of the function,

   $\sigma^2 = \beta_1 + \beta_2 U^K$

   where exponent K is user defined, returned a spurious "failure" error message in some cases and some contexts. When this function specific problem did occur it was limited to manual estimation or estimation as part of an Auto-processing run, but it did not occur in the multi-estimation context. Now corrected.

4. When blank specimens are specified during data importing the program assumes that the intention is Limit of Blank (LoB) and Limit of Detection (LoD) calculations and, if necessary, automatically shifts all values up by the smallest amount necessary to ensure all mean values are > 0. This in turn ensures the availability of all variance functions in LoB, LoD calculations (LoB, LoD estimates are back-shifted appropriately). While all mean values are > 0 in these cases there will inevitably be some negative *replicate* values. The constant CV function,

   $\sigma^2 = \beta_1 U^2$

   can be legitimately estimated (since all mean values are > 0), but the program crashed if an attempt was made to variance stabilise the data (logarithm of a negative number). The variance stabilisation option should have been suppressed in this specific situation. Now corrected. Note that variance stabilisation expressions derived from other variance functions remain legitimate in these cases.

Consult on-line help topic Contents > Appendix for details of fixes across all program versions.

References

1. Sadler WA, Smith MH. Applications of a three-parameter variance function in statistical analysis of immunoassay data. In: Developments in Radioimmunoassay and Related Procedures (Proc IAEA Symp., Vienna 1991), IAEA, Vienna, pp 383-392 (1992)

2. Ekins RP. The precision profile: its use in assay design, assessment and quality control. In: Hunter WM, Corrie JET, eds., Immunoassays for clinical chemistry, 2nd ed., Edinburgh: Churchill Livingston, pp 76-105 (1983)

3. Sadler WA, Smith MH. Estimation of imprecision in immunoassay quality assessment programmes. Ann Clin Biochem, 24:98-102 (1987)

4. Sadler WA, Smith MH, Legge HM. A method for direct estimation of imprecision profiles, with reference to immunoassay data. Clin Chem, 34:1058-61 (1988)

5. Sadler WA, Smith MH. A reliable method of estimating the variance function in immunoassay. Comp Stat & Data Anal, 3:227-239 (1986)

6. Tate J, Koerbin G, Potter J, Hickman P. Evaluation of the Beckman Coulter IMMAGE high sensitivity C-reactive protein (CRPH) assay. Clin Chem Lab Med, 40(Suppl):S226 (2002)

7. Sadler WA. Error models for immunoassays. Ann Clin Biochem, 45:481-485 (2008)

8. Daniels PB. The fitting, acceptance, and processing of standard curve data in automated immunoassay systems as exemplified by the Serono SR1 analyzer. Clin Chem, 40:513-517 (1994)

9. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet, i:307-310 (1986)

10. Rocke DM, Lorenzato S. A two-component model for measurement error in analytical chemistry. Technometrics, 37:176-184 (1995)

11. Bartlett MS. The use of transformations. Biometrics, 3:39-52 (1947)

12. www.westgard.com/guest12.htm (C.G. Fraser)

13. www.westgard.com/guest7.htm (P.H. Petersen)

14. www.westgard.com/biodatabase1.htm (Biological variation database)

15. Sadler WA. Joint parameter confidence regions improve the power of parametric regression in method-comparison studies. Accred Qual Assur, 15:547-554 (2010)

16. Deming WE. Statistical adjustment of data, New York: John Wiley & Sons (1943). Reprinted, New York: Dover Publications Inc., pp 184 (1964)

17. ISO 11843-1:1997. Capability of detection – Part 1: terms and definitions.

18. Sadler WA. Using the variance function to estimate limit of blank, limit of detection and their confidence intervals. Ann Clin Biochem, 53:141-149 (2016).

19. Sadler WA. Variance functions, detectability and bias: a re-evaluation. Ann Clin Biochem, 54:286-288 (2017)

20. Hawkins DM. Diagnostics for conformity of paired quantitative measurements. Stat in Medicine, 21:1913-1935 (2002)

21. Hawkins DM. A general variance model in methods comparison. J Chemometrics, 27:414-419 (2013)

22. Sadler WA. Using the variance function to generalize Bland-Altman analysis. Ann Clin Biochem, 56:198-203 (2019)

23 Breusch TS, Pagan AR. A simple test for heteroscedasticity and random coefficient variation. Econometrica, 47:1287-94 (1979)

24. Cook RD, Weisberg S. Diagnostics for heteroscedasticity in regression. Biometrika, 70:1–10 (1983)

25. Sadler WA, Smith MH. A computer program for estimating imprecision characteristics of immunoassays. Comput Biomed Res, 23:105-114 (1990)

26. Sadler WA. A new Win32 computer program for estimating immunoassay variance functions. Computer Meth & Prog in Biomed, 67:195-199 (2002)

27. Sadler WA. An upgraded Win32 computer program for estimating immunoassay variance functions. Clin Chem Lab Med, 39(Suppl):S133 (2001)