LinRegPCR (11.0)

Analysis of quantitative RT-PCR data

J. M. Ruijter
S. van der Velden
A. Ilgun

Heart Failure Research Center
Academic Medical Center
Amsterdam, the Netherlands

CONTENTS

1. About this manual

The first two chapters of this manual contain the instructions on the basic use of the LinRegPCR program, starting with the installation of the program and importing data from the qPCR system into Excel. Detailed descriptions of the interactive procedures of LinRegPCR are given in the chapter 3 of this manual. Chapters 4 and 5 describe the menus and different Tab pages, respectively. The LinRegPCR web page (http://LinRegPCR.nl) gives answers to some frequently asked questions.

1.1 Disclaimer

LinRegPCR is based on the procedures described in the paper by Ruijter et al., Nucleic Acids Research, 2009. The procedures in this program are to be used for fluorescence data resulting from monitoring the PCR reaction with an intercalating dye like SYBR Green or a related chemistry. They are not valid for hybridisation probes that become fluorescent during the elongation by polymerase and of which the fluorescence accumulates.
By opening and using this software you acknowledge that you have read the above paper, understand it, and agree with its conclusions. Therefore, you assume all responsibility and liability for the selection of this program to achieve your intended goals, and for the conclusions you draw from the results. The authors cannot be hold responsible for any consequences of the use of this program.

1.2 Update history of LinRegPCR

Version 11.0 (24-01-2009)**:**
Version number is set to 11 because of reference in Ruijter et al., Nucleic Acids Research 2009.

- Added input format for Stratagene systems (Format 1, Vertically Grouped data)
- Added export of input data in '1 row per sample, 1 column per cycle' format as a kind of 'transpose data' option for Stratagene and MJ Research users

Version 10.3:

- Implemented detection of noisy datasets (see 3.10).
- Removed error that lead to crash in the setting of the window-of-linearity in noisy datasets.
- Extended quality check with 'noisy sample' (see 3.11).

Version 10.2:

- Corrected error in the setting of the Ct value.
- Corrected assignment of upper W-o-L limit for first amplicon group.
- Implemented full reset of the values when 'determine baseline' is pressed twice.

Version 10.1:

- 'division by zero' error that occurred when not enough data were present at low W-o-L settings was captured and handled.
- Freezing of the tab pages when a grouping error occurred was removed.

Version 10 (22-03-2008)**:**
In this version the following features have been implemented:

- import of raw (not baseline-corrected) data files
- estimation of baseline fluorescence
- new approaches to determine the window-of-linearity and fluorescence threshold

- definition of sample groups per amplicon
- calculation of the mean PCR efficiency per amplicon group
- calculation of starting concentrations based on mean efficiency and individual Ct values
- reports on data quality for each sample

Non-released versions 8 and 9 were intermediate versions while implementing baseline subtraction and amplicon grouping.

Distribution of LinRegPCR version 7.4 (February 2004) is discontinued.

A recent update history can be found at: http:/LinRegPCR.nl . Please register yourself as LinRegPCR user on this website. When you do, you enable us to send you information on updates and bug fixes.

1.3 Known bugs

Opening a second data set after analyzing the first sometimes gives an error, especially when the second set is smaller than the first. Therefore, it is better to close the program and start again when you have more data sets to analyze.

1.4 Installation and requirements of LinRegPCR

The version of LinRegPCR that you received has been zipped. In the LinRegPCR.zip file you will find:
The executable program: LinRegPCR.exe
This help file in Adobe Acrobat readable format: LinRegPCR_help.pdf
Unzip the files into a directory and you are ready to run. No special configuration is needed. For easy use it is handy to put shortcuts to the program and the help text in your Start menu or on your Desktop.

LinRegPCR stores some of the choices you make during the use of the program in the Windows Registry (Key: HKEY_CURRENT_USER\Software\AEL Research\LinRegPCR). This enables the program to select these choices as default the next time you use the program.

The program requires Windows 95 or later and Microsoft Excel 97 or later. The LinRegPCR program expects Excel to use the decimal separator that is defined in Windows. To make sure of this, in Excel go to ***Tools - Options*** and in the ***Options*** dialog, on the ***International*** Tab, check ***Use system separator***.

1.5 References

The basics of this program are described in:

Ruijter JM, Ramakers C, Hoogaars W, Bakker O, van den Hoff MJB, Karlen Y, Moorman AFM. (2009) Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Research, in press.

The calculation of efficiencies for individual samples from a subset of data points of the amplification curves is explained in:

Ramakers C, Ruijter JM, Lekanne Deprez RH, Moorman AFM. (2003) Assumption-free analysis of quantitative real-time PCR data. Neurosci Letters 339: 62-66, 2003.

The linear regression procedure can be found in every statistics textbook. **Windows** and **Excel** are trademarks of Microsoft Inc.

1.6 Contacts

Frequently asked questions: http://LinRegPCR.nl

Please direct questions and remarks to info@LinRegPCR.nl . This mail will be checked once per week.

2. Introduction to the LinRegPCR program

LinRegPCR is a program for the analysis of real time RT-PCR Data, also called quantitative PCR (qPCR) data based on SYBR Green or similar chemistry. The program uses non-baseline corrected data, performs a baseline correction on each sample separately, determines a window-of-linearity and then uses linear regression analysis to fit a straight line through the PCR data set. From the slope of this line the PCR efficiency of each individual sample is calculated. The mean PCR efficiency per amplicon and the Ct value per sample are used to calculate a starting concentration per sample, expressed in arbitrary fluorescence units. Data input and output are through an Excel spreadsheet.

2.1 Definition of terms used in this manual and in the program

Amplicon group: A set of samples in which the same pair of primers is used to amplify the DNA-of-interest

Background: Fluorescence of the reference fluorochrome (ROX or Fluoresceine) used to correct for variation in pipetting and/or fluorescence outside the reaction wells. The background is handled by the PCR system.

Baseline: Measured fluorescence when no amplification-specific fluorescence can yet be determined. This fluorescence includes fluorescence from cDNA, primers and unbound SYBR Green.

Ct: Number of cycles needed to reach the fluorescence threshold. Samples with higher starting concentrations will reach this threshold earlier and will have a low Ct value; Ct is used to calculate N0.

E or PCR Efficiency: Amplification efficiency expressed as a value between 1 and 2. This efficiency is calculated from slope of the amplification curve in the exponential phase. Ideally the PCR efficiency is 100%, meaning that in each cycle the amount of amplicon doubles (E=2). A value of 1 means no amplification.

Exponential or log-linear phase: The section of the PCR curve which best represents the exponential phase of the PCR reaction, when the levels of generated fluorescence exceed baseline fluorescence, but reagents have not yet begun to be limiting. In this phase the amplification efficiency is similar across samples regardless of the starting concentration.

Fluorescence threshold or **Nt**: Fixed amount of fluorescence, used to determine the Ct value. By default set at one cycle below the upper limit of the window-of-linearity.

N0: Starting concentration per sample, expressed in arbitrary fluorescence units.

Plateau phase: Endpoint of the PCR reaction - the phase in which there is significant depletion of one or more reaction components. In the plateau phase the amplification curves of the real-time PCR are no longer exponential and the PCR efficiency dropt to zero.

Raw data: Fluorescence values that are background-corrected but not baseline-corrected by the PCR system.

Window-of-Linearity or **W-o-L**: Subset of data points in the exponential phase that is used to determine the PCR amplification efficiency (E) per sample. The Window-of-Linearity is set automatically by the program but can be moved by the user. Default the W-o-L width is set to 4 data points or amplification cycles.

2.2 Basic procedure for using the program

1. Export background- but not baseline-corrected data from the qPCR system to Excel
2. Read these raw data from Excel into LinRegPCR

You can check the ***Data*** page; sub-page ***Input***, to see whether all data are there.

3. Press ***Determine Baselines***

The program will do a baseline correction per sample and will continue with determining a common window-of-linearity for all samples. The program then sets the fluorescence threshold at one cycle below the upper limit of this window.

4. Define grouping of samples per amplicon
   if needed: manually edit the amplicon group assignment.

The program sets a window-of-linearity per amplicon group and sets the fluorescence threshold for each group.

5. Check the individual samples and if needed:
   - correct the baseline in individual samples.
   - adjust the window-of-linearity for an amplicon group or an individual sample.
   - export graphs to clipboard (click the right mouse button on a graph).
6. Save results to Excel

2.3 Exporting raw data from the PCR apparatus

The LinRegPCR program reads data from an Excel spreadsheet. Therefore, the data have to be exported from your PCR apparatus and imported into Excel. All real-time PCR apparatus enable the export of the fluorescence data per cycle into a text file format.

Note that the exported data have to be corrected for the background but should NOT be corrected for the fluorescence baseline. LinRegPCR estimates this baseline per sample and does a baseline subtraction.

Please find out for your own PCR apparatus how to export raw data. You are on your own at this point. There are too many different types of machines for us to keep track of all these procedures. Sorry.

One tip may help: When you cannot find the export of raw data, try to set the baseline to 0 or to the minimum observed value per sample.

2.4 Importing data into Excel

Most PCR systems export data to text files. A tab-delimited (*.TXT or *.DEL) or a comma-delimited (*.CSV) text file can be opened in Excel:

1. Start Excel.
2. Drag the file from the Windows Explorer onto the running instance of Excel

Or:

3. Drag the file onto the Excel icon in the Windows status bar and wait for Excel to maximize.

Then:

4. Drop the file onto the open worksheet.

In most cases Excel will directly recognise the structure of the file and split the file over the appropriate number of columns. If not, each row of data will be imported as a long line of text in the cells in column A. To split these lines into columns:

1. Click on the grey A to select the whole column A
2. Select ***Data – Text to columns*** from the main menu
3. Check ‘***Delimited’*** and press ***Next***
4. Check ***‘Tab’*** or ***‘Comma’*** and press ***Finish***

You can also open a text file in Excel:
1. Select ***File*** – ***Open*** from the menu
2. In the Open-dialog set file type to '***Text files'***
3. Choose the file you want to open
4. In the dialog check ***'Delimited'*** and press ***Next***
5. Check ***'Tab'*** or ***'Comma'*** and press ***Finish***

You have to save the imported data as an .XLS file before you can continue with the analysis.

Trouble shooting:

- Make sure Excel is in ***'Ready'*** mode (in the left bottom corner of Excel). When you are editing a cell that has to be read by LinRegPCR, importing the data into LinRegPCR will fail. LinRegPCR will then display the message "Call was rejected by callee ".

LightCycler 480 users will find that the raw data output results in a text file of one line per cycle and melting curve point per sample. To convert the amplification data in this text file into an Excel sheet that can be read by LinRegPCR they can use the program LC480Converter that can be obtained by clicking this link: http://HFRC.nl , got to ***Downloads*** and find the **LC480 Conversion** program under ***Applications***.

Please register as user of LC480Converter by clicking the button in the ***About*** box of this conversion program.

3. PCR data analysis with LinRegPCR

3.1 Reading data into LinRegPCR

Make sure the Excel file containing the data is open!

1. Start LinRegPCR. The program will immediately open the ***Read-from-Excel*** dialog (Figure 1). The active Excel book and sheet will be shown in the ***book*** and ***sheet*** list boxes. If these are not the book and sheet that you need:
2. Choose the book and sheet from the dropdown lists that will appear when the arrow head next to the boxes is pressed.
   If the book or the sheet does not show up in the list boxes:
   a. Press ***Cancel*** and go to Excel
   b. Save your data file as an Excel workbook (*.XLS)
   c. Go back to LinRegPCR and choose ***File – Read from Excel*** from the menu
   d. Return to step 2.
3. Select the ***data file format*** by clicking the appropriate radio button.
4. Set the ***column*** letters and ***row*** numbers to define the range of cells that has to be read. The format you choose and the column and row values you give will be saved in the Windows Registry. The next time you use the program these values will be displayed as default.
5. Set baseline-corrected ***Yes*** or ***No***. The program expects data that are not baseline-corrected. When your data are baseline corrected by the PCR system you have to indicate this in the radio button choice in the ***Read-from-Excel*** dialog. When you forget to do this, you have to use the menu option ***Baseline - Set Baseline Corrected*** to skip the baseline correction of LinRegPCR.
6. Press ***OK***.

Do not close Excel!

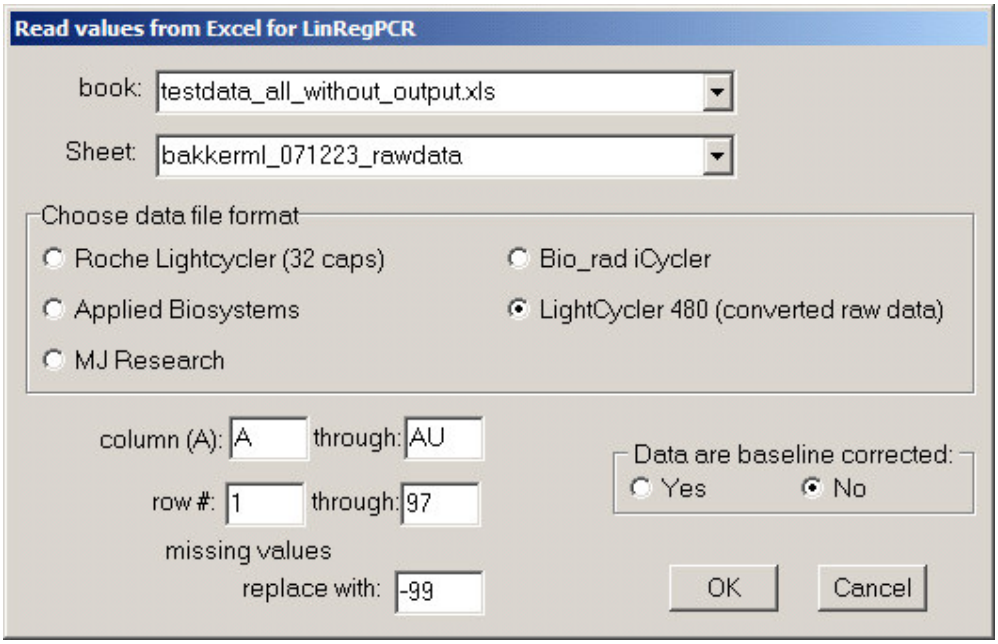

Figure 1: Read from Excel dialog. You have to choose the workbook and worksheet with your data (the currently open workbook and sheet are displayed by default), the format of your input file and the range of cells that you want to read. By default the program expects the data to be not baseline-corrected. If they are, you have to choose the ***Yes*** radio button.

LinRegPCR recognizes five input formats (Fig. 1). For each format it expects a fixed number of leading columns and/or header rows. When your PCR apparatus is not in this list, it will almost certainly have an output format that fits with one of those in the list. If not, try to add rows and columns to make it fit, or try to merge columns to remove excess columns. If you cannot get it right, contact us and we will try to find a solution.

Implemented export formats :

Roche LigthCycler (32 caps):

This format consists of 10 header rows that are ignored by LinRegPCR and two columns per sample starting with a two row header followed by cycle number - fluorescence value pairs. The range you have to give in the columns and rows edit fields is inclusive the two header rows per sample. For a 40 cycle PCR and the full 32 samples the range would be: columns A thru BL and rows 11 thru 52. The entries in the left cell of first header row per sample will be used as sample name; the right cell and the second header row will be ignored. The numbers in the first column per sample are used as cycle numbers.

Applied Biosystems

This format consists of one header row (with cycle numbers) and two leading columns around a data block that consists of one sample per row, one cycle per column. The whole range has to be given in the columns and rows edit fields. For a 40 cycle PCR and 72 samples the range would be: columns A thru AP and rows 1 thru 73. The two leading columns will be combined to one sample name, the entries of the columns being separated by a space. The header row will be converted into integer cycle numbers.

MJ Research

This format consists of one header row (with sample identification) and three leading columns around a data block that consists of one sample per column and one cycle per row. For a 40 cycle PCR and 60 samples the range would be: columns A thru BK and rows 1 thru 41. The entries in the header row are used as sample name. The entries in three leading columns are ignored and the row number in the data block is used as integer cycle number.

Bio-Rad iCycler

This format consists of two header rows and one leading column around a block of data that consists of one sample per column and one cycle per row. The leading column contains fractional cycle numbers, the first header row is empty except for the word ‘well’ in cell A1 and the second header row contains the well identification. Both header rows have to be included in the range given in the columns and rows edit fields. For a 45 cycle PCR and 30 samples the range would be: columns A thru AE and rows 1 thru 47. The entries in first header row are ignored by the program; those in the second row are used as sample names. The leading column is converted into fractional cycle numbers.

Roche LightCycler 480 (converted raw data: see http://LC480Converter.hfrc.nl)

This format is the same as the Applied Biosystems format (see above)

Stratagene Format 1 (Vertically Grouped by Sample)

This format consists of two header rows ('Amplification Plots' and an empty row) followed by a data block per sample This block starts with a sample header row with the entries sample name, Cycles and Fluorescence (Rn) followed by one row per cycle. The total input block is therefore 2 + Nsamples * (Ncycles +1) rows long and 3 columns wide. For a 40 cycle PCR with 66 samples the range would be: columns A thru C and rows 1 thru 2708 {=2 + 66 *(40 +1)}. The two header rows are ignored, the entry in column A of the sample header row is used as sample name and the entries in column B are used as integer cycle numbers.

NOTE: Although some rows and columns are ignored by the program, they have to be given in the range you enter in the columns and rows edit fields of the ***Read-from-Excel*** dialog. Also note that these formats may already be different from the ones that are generated by the current versions of the software of the PCR apparatus. In that case, try to find a format that fits to the exported data or contact us.

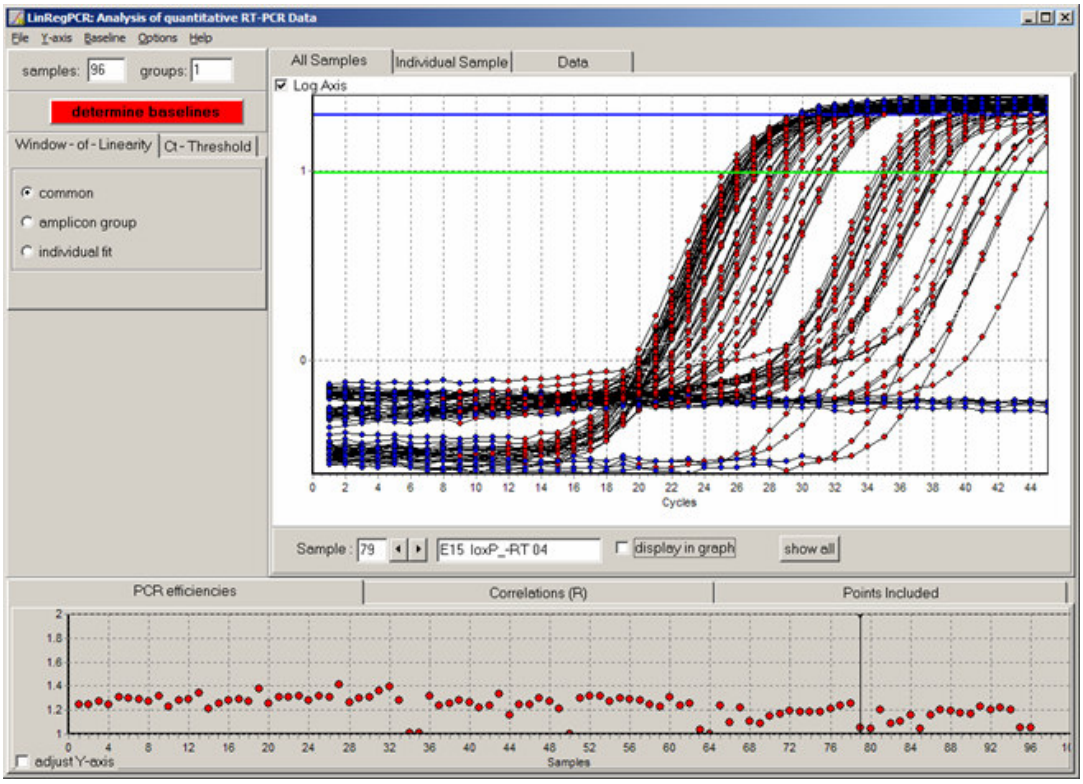

Figure 2: LinRegPCR immediately after importing a not baseline-corrected data set. Note that the ***Determine Baselines*** button is highlighted to indicate the next step of the analysis.

Trouble shooting:

- Excel has to be running, and the book (spreadsheet) with your data should be opened.
- Excel should NOT be in edit mode: the status bar of Excel has to display **Ready**. To make sure of this: place the cursor outside the data range and press ***Enter***.
- Some sheet names are not compatible with LinRegPCR. When a sheet does not turn up in the drop down list: cancel the ***Read-from-Excel*** dialog, go to Excel, rename the sheet and try again.
- Excel Chart sheets are incompatible with LinRegPCR.

- LinRegPCR expects that Excel uses the same decimal separator as Windows. To make sure this is the case: in Excel go to ***Tools - Options***, choose ***International*** and set Excel to ***Use system separators***.
- Sometimes the book or sheet does not show up in the list boxes. In that case check for typing errors in the file extension (*.TXT or *.CSV). You can almost always solve this problem by saving your data file as an Excel-workbook (*.XLS).
- Note that Excel may have 'hidden' sheets that it uses for storage of macros and functions. These sheets may turn up in the sheets-listbox as names that you do not recognize as your data sheets. NEVER choose one of those sheets.

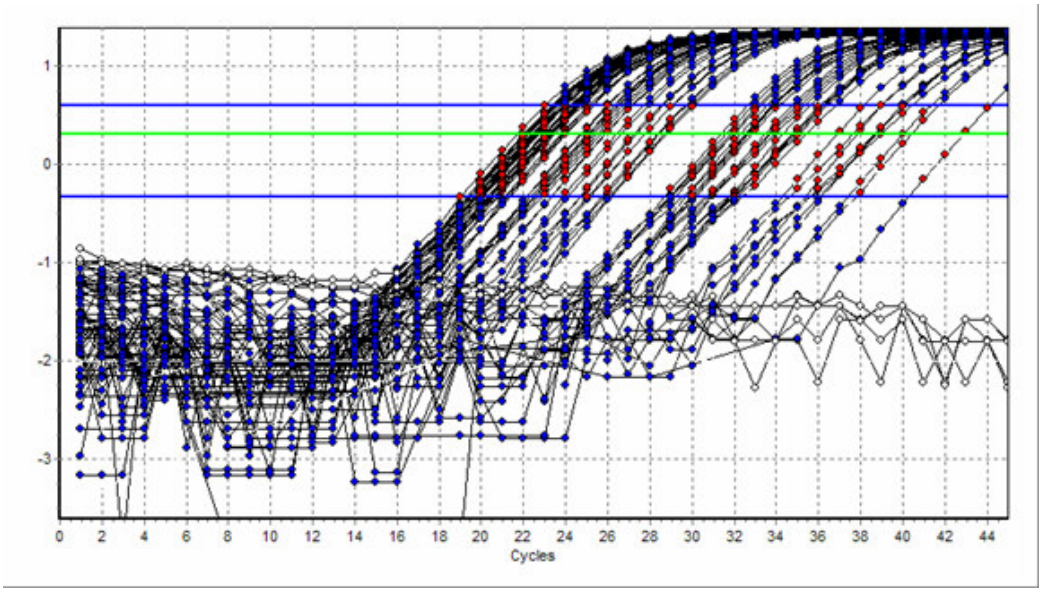

Figure 3: All Samples graph after estimation and subtraction of the baseline.

3.2 Baseline Estimation

LinRegPCR determines the baseline fluorescence **per sample** by reconstructing the log-linear phase. It determines the cycle at the maximum of the Second Derivative (SDM) of the observed fluorescence values in each sample which is the start of the plateau phase. Then it does an iterative search for a baseline value that gives the longest straight line of data points down from the plateau phase. During this process it determines the following 'deviating' samples:

- **no-amplification**: samples with less than 7-times overall increase. This assignment can be overruled by un-checking the ***Mark sample as excluded*** checkbox in the lower right of the individual sample window. You can also use this check box to exclude samples that you do not want to include in the calculation; they will be marked as **User Excluded** in the results.
- **baseline error**: samples for which no baseline value can be determined because no fluorescence value between the SDM and the minimum observed value results in a straight line. This assignment can be overruled by setting a manual baseline. The menu option ***Baseline*** - ***Manual Baseline*** displays the manual baseline controls (see 3.6).
- **noisy sample**: sample in which the fluorescence in cycle x is higher than in cycle x+1, in a W-o-L just below the SDM cycle. This high in the exponential phase, the fluorescence values should be continuously increasing.

- **no-plateau**: samples that do not reach the plateau phase. This assignment cannot be overruled.

These samples are flagged and can be separately displayed using the menu item: ***Options - Display Flagged subsets***. The **no-amplification** and **baseline error** samples will be ignored by the program. They are always excluded from the calculation of the mean PCR efficiency and will receive a starting concentration of -999.

The user can decide to include or exclude the **no-plateau** samples from the calculation of the mean PCR efficiency (see 3.5). For these samples a starting concentration will be calculated when they are not also an **efficiency outlier**.
For details on the baseline estimation procedure: see Ruijter et al. , Nucleic Acids Research, 2009.

In some datasets, the baseline estimation will result in a warning that the dataset is **too noisy** to be analyzed automatically. This does not have to mean that the dataset is completely lost. The program will then display the manual baseline and window-of-linearity controls. See 3.8 for help on the handling of noisy datasets.

3.3 Amplicon groups

The subset of samples in which the same pair of primers is used is called an amplicon group. Because the PCR efficiency per amplicon is assumed to be constant, but individual samples have slightly variable observed PCR efficiencies, LinRegPCR uses the mean PCR efficiency per amplicon in its calculations. To allow this, the samples have to be assigned to amplicon groups. The Amplicon Groups page (Fig. 4) gives three choices:

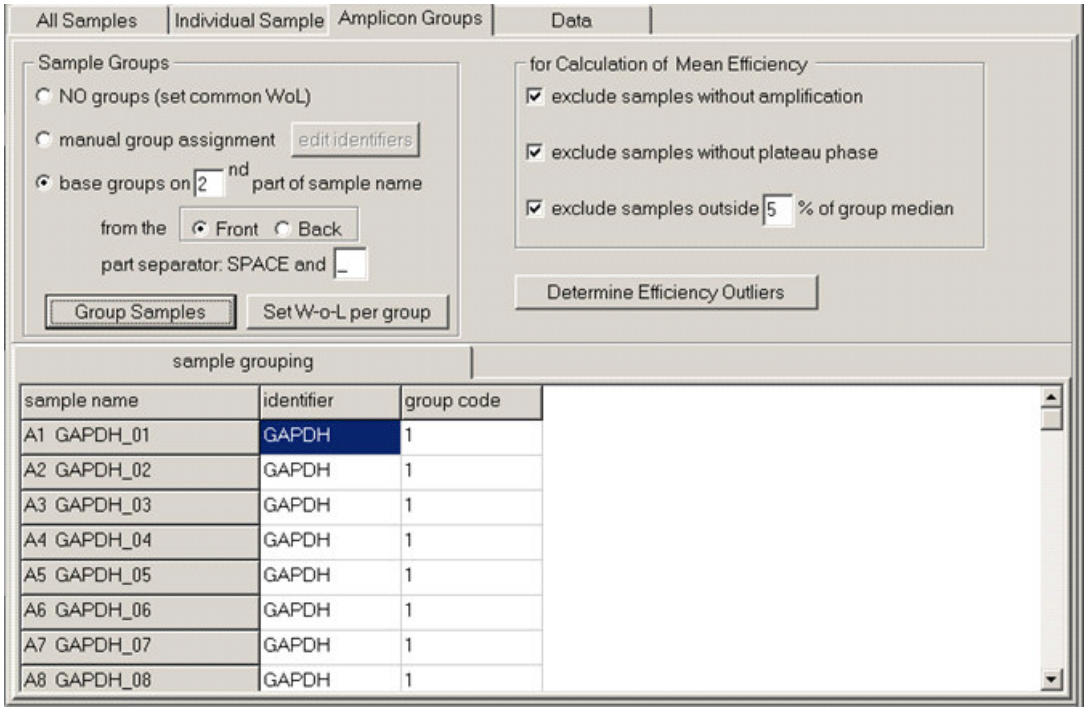

Figure 4: Amplicon Groups Tab after setting the grouping parameters to 'base groups on the second part from the front of the sample name using space and underscore as separators'.

- **No groups**: all samples are one group.

When this is the case you do not have to do anything. After the baseline correction a common window-of-linearity is set assuming that in all samples the same amplicon is produced.

- **manual group assignment:**
  When the sample names do not contain information on the amplicon, you will have to assign the grouping manually.
  1. choose the radio button ***'manual group assignment'***
  2. press the ***edit identifiers*** button

  This button puts the sample group grid in edit mode. You can now enter values into the **identifier** column of this grid. Activate a cell by double clicking and then enter the identifier. A group code is assigned automatically.
  When you have given all identifiers:
  3. Press the ***apply groups*** button (which is the same button you just pressed).

  The program now sets a window-of-linearity and fluorescence threshold per amplicon group.

- **base groups on part X of the sample name** (see Figure 5)
  When the sample name contains the amplicon identifier in a fixed format, you can base the group assignment on the sample name.
  1. check the radio button 'base groups on etc'
  2. enter the index of the part that describes the amplicon
  3. indicate whether you are counting from the front or the back of the sample name
  4. when the separator is not (only) a space, give the extra separator
  5. press 'Group Samples'

  The program now extracts the amplicon identifier and assigns group codes. Check these group assignments and change the above entries when required. If you want you can switch to 'manual group assignment' to correct the identifiers.
  When the assignments are correct:
  6. press 'Set W-o-L per Group'

  The program now sets a window-of-linearity and fluorescence threshold per amplicon group.

You can look at the groups separately by browsing through the groups with the controls that have appeared on the lower left of the screen. You can select individual samples by clicking on the sample in the PCR efficiencies graph.

***NOTES*:**
- LinRegPCR has now set a window-of-linearity per sample. When you browse through the samples you can now also make corrections in this W-o-L setting (see 3.7)
- Naming samples: amplicon group assignment based on the sample name is recommended. To do this it is easiest to enter the sample names in Excel. It is best to also keep the well code (e.g. A1) in the sample name because this is the link to your pipetting scheme. The different parts of the sample name can be separated by spaces or by another character. But make sure the separator is unique and not a part of your amplicon identifier. For this reason, the identifier cannot contain spaces. With a little trial and error your will soon find your preferred format.
- Minus-RT samples. In defining amplicon groups take care of handling the minus-RT samples. When you are sure the amplified product in these samples is your amplicon-of-interest, you can include them in the amplicon group. However, when the product may be

something different (e.g. genomic DNA in stead of cDNA) the minus-RT samples may display a different PCR efficiency. In such a case it is better to treat them as a separate amplicon group.

3.4 Check individual samples

The baseline estimation cannot be completely fail-proof. Therefore, the user of LinRegPCR is urged to browse through the individual samples to check deviating samples that were missed by the program. The user can exclude a sample by checking the box at the lower right next to the graph. The user can also use the manual baseline controls to try to correct the baseline in a sample when the program failed to find the right baseline because the data were too noisy (see 3.6 and 3.8). When samples are marked as excluded, or when the baseline of a sample is manually corrected, the user should press '***Recalculate W-o-L***' to adjust the W-o-L to the new data.

When required, the user can also change the settings of the common or amplicon group window-of-linearity at this stage. For individual samples, data points can be included in - or excluded from - the W-o-L. This will set an individual W-o-L for that sample. Common or group W-o-L setting can be changed by manual W-o-L settings (see 3.7).

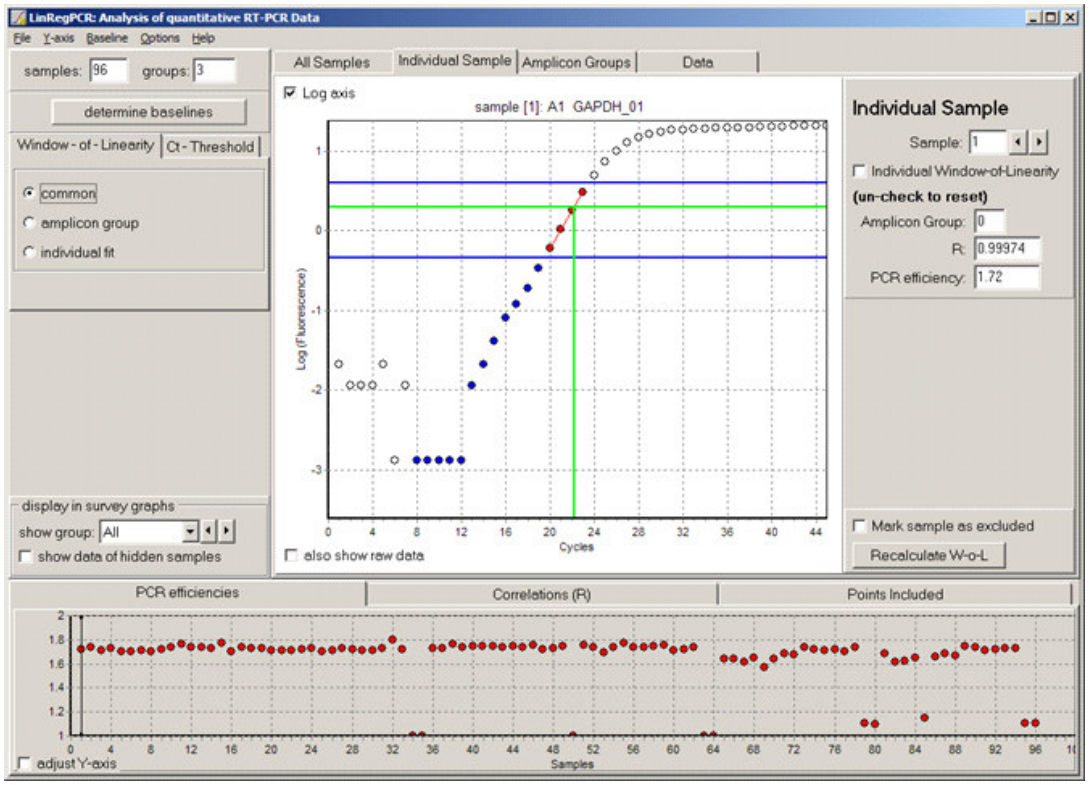

Figure 5: LinRegPCR showing the individual samples to enable the user to browse through the samples and check the baseline correction. Samples with strongly deviating PCR efficiency, as shown in the bottom panel, require special attention.

3.5 Mean efficiency

The group of samples in which the same pair of primers is used are called an amplicon group. PCR efficiency per amplicon is assumed to be constant; nevertheless individual samples have slightly variable observed PCR efficiencies. Therefore, it is recommended to use the average efficiency of all samples for each amplicon group to do the quantification. For this reason LinRegPCR uses the mean PCR efficiency per amplicon in its calculations (see Ruijter et al. , Nucleic Acids Research, 2009.).

Amplicon groups can be assigned with the controls on the Amplicon Groups page (see 3.3). On the top right of this page you can also make some decisions on the inclusion or exclusion of deviating samples from the calculation of the mean efficiency.

- Samples without amplification are always excluded.
- Samples that do not reach the plateau phase can be excluded. Excluding these samples may give you a better mean efficiency but, because this mean is then based on fewer samples, it may not be the best mean.
- Samples with a deviating individual PCR efficiency can be excluded from the calculation of the mean efficiency. You can give the range around the median efficiency that you want to allow. The default range of 5% means that around a median efficiency of 1.83, a range from 1.78 to 1.88 is allowed. This exclusion will only affect the mean efficiency when the distribution of observed individual efficiencies is much skewed.

When you have chosen to exclude efficiency outliers, you can display those samples by choosing the menu option ***Options - Display Flagged Subsets*** and then choose the set from the list box that appeared at the left side of the window.

***NOTES*:**

- Excluded samples are only excluded from the calculation of the mean efficiency per amplicon group. The sample will receive a starting concentration (N0) in the output.
- The choices you make will be saved to the Windows Registry and will be loaded as default the next time you run the program.

3.6 Fluorescence baseline

The fluorescence baseline is determined as described in chapter 3.2. In some cases the program will not be able to set the correct baseline. This is the case when the baseline-to-plateau ratio is too low or when the observation noise is too large. Both problems make that the data points downwards from the plateau phase do not form a continuously decreasing set of points. In that case the program restricts the log-linear phase to the small number of points that are continuously decreasing from the plateau phase and the baseline estimate will be wrong, in most cases too high. Manual baseline setting might rescue those data. Also see 3.8 for handling of noisy data. Another reason to use the manual baseline setting is to rescue samples that were assigned a baseline error.

Manual setting of the fluorescence baseline

When the ‘also show raw data’ check box on the ‘Individual Sample’ graph is checked, the not baseline-corrected data are shown as red points and the baseline is indicated by a red line. The user can change the baseline by drag-and-drop of the red baseline. When the user does so, the manual baseline controls will appear at the right bottom of the ‘Individual sample’ page.

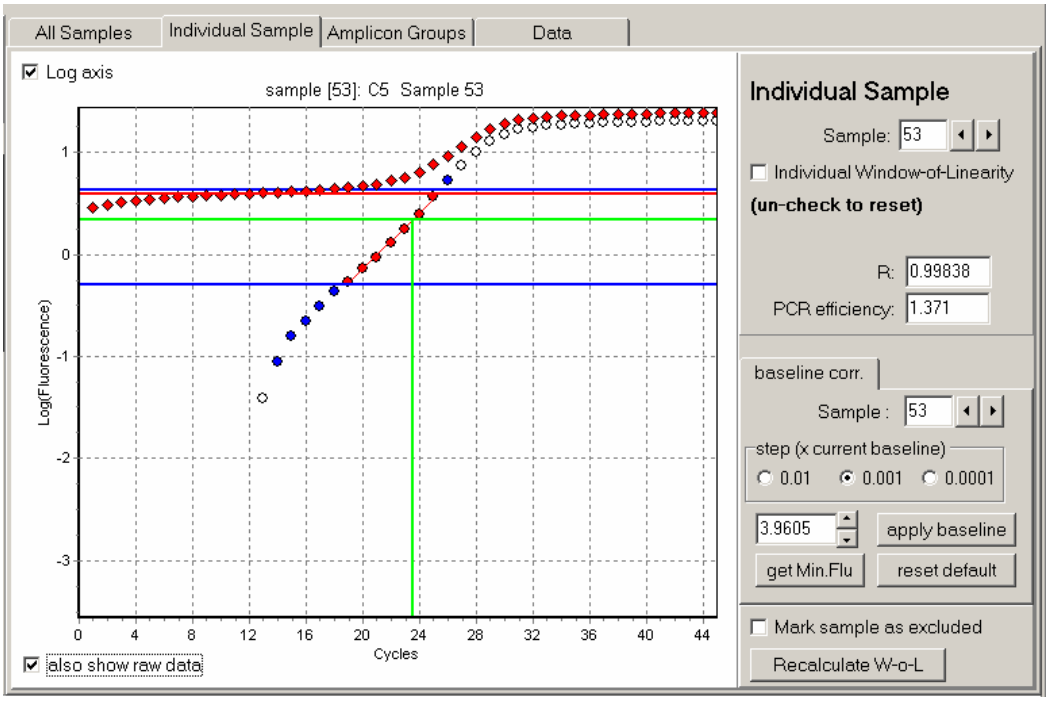

Figure 6: Individual Sample page showing the baseline-corrected sample (white, blue and orange points), the raw data (red points) and the fluorescence baseline (red line). The baseline can be estimated manually with the controls on the lower right of the screen.

The baseline value can be entered in the edit field or changed stepwise by clicking the up-down arrows. The step size is determined by the choices (0.0001 thru 0.01 of the current baseline). When setting a manual baseline is required, the aim should be to construct a straight line of data points downwards from the plateau phase.

When you start to manual threshold setting is can be a good idea to start with the minimum observed fluorescence of the sample:
- press the '***get Min.Flu.***' button to have the value displayed in the edit field
- press '***apply baseline***' to subtract this value from the raw data

***NOTE*:**
- When the current baseline is 0 (=zero) the up-down buttons will have no effect.
- When a manual baseline is set, the window-of-linearity should be recalculated by pressing the ***Recalculate W-o-L*** button at the bottom right of the page.
- When you apply a manual baseline to a 'baseline error' sample, the sample will be included in the calculations, unless you check the '***Mark sample as excluded***' checkbox.
- Samples with a manual baseline will be labeled as such in the output (see 3.11)

3.7 Window-of-Linearity and Nt threshold

In this version of LinRegPCR you can distinguish between three window-of-linearity settings:
- **common window**, which you should use when all samples are amplified with the same pair of primers.
- **amplicon group window**, which you should use when different sets of primers are used (see 3.3 and 3.5).
- **individual window**. This option is included to provide of backward compatibility. When you choose this option the program will set a W-o-L per sample in the same way it did in

LinRegPCR; version 7.4. You will see that the variation in PCR efficiency values will increase significantly.

After the baseline correction a common W-o-L is automatically set. After the definition of amplicon groups a W-o-L per amplicon group is automatically set. All W-o-L's can be adjusted manually using the W-o-L controls that are displayed when you choose the menu option ***Display W-o-L Controls*** from the ***Options*** menu (see chapter 4).

The Nt threshold is automatically adapted to fit the window-of-linearity choice.
- The common threshold is used when you choose a common window or when you have chosen to use individual windows.
- The threshold per amplicon group is used when you choose amplicon windows.

***NOTES*:**
- In case you want to do your own comparative Ct calculations with the exported efficiency, threshold and Ct values it is required to use a common threshold for all amplicon groups.
- You can set a manual threshold using the edit field on the Ct Tab of the notebook on the top left of the screen. When you have defined amplicon groups, the given threshold will be set for the currently displayed group.

The N0 results that LinRegPCR reports are NOT affected by the choice of Nt threshold. However, obviously the Ct values do depend on the threshold.

Manual setting of the Window-of-Linearity
The blue lines in the ***All Samples*** graph and the ***Individual Sample*** graphs indicate the upper and lower limit of the window-of-linearity. Data points between these lines are used by the LinRegPCR to calculate the PCR efficiency per sample. The user can change these window settings manually if required.

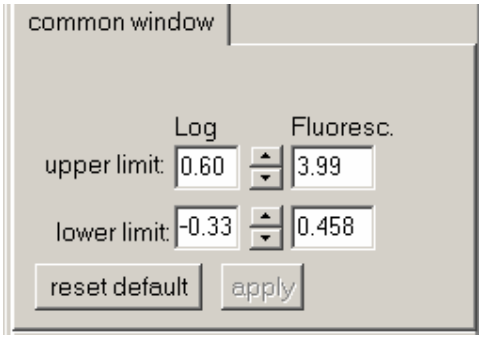

Figure 7: Controls that can be used to set a common window-of-linearity. Data can be entered in the edit fields, or the up-down arrows can be clicked. When the edit fields are used, the values are applied after pressing the ***apply*** button.

Common window.
When a common window is used the user can drag-and-drop the window limits in the ***All Samples*** graph. Click on the blue lines until a vertical ↔ cursor appears, drag the line and click again to drop the line. This action will also open the ***Common Window*** controls (middle left on the screen) that allow you to use the up and down arrows to fine tune window settings. You can also fill in values that you know from earlier analyses.

Amplicon windows.
When amplicon groups have been defined, and you are browsing through the amplicon groups, the ***All Samples*** graph will show the window-of-linearity per amplicon group. You can drag-and-drop the window limits or use the amplicon window controls on the middle left of the screen to adjust the window.

Individual windows.
When you drag-and-drop the blue window-of-linearity limits in the ***Individual Sample*** graph, the sample will be marked as having an individual window. This will also display the individual window controls at the middle right of the screen. You can also include or exclude a data point in the individual window by clicking on the point itself.

***NOTES*:**
- The common and amplicon group windows set automatically by LinRegPCR are the windows that give the least variation between individual efficiency values per group. Adjusting the window manually will lead to increased variation between samples.
- When an individual window is set, the resulting PCR efficiency will be used in the calculation of the mean efficiency for all samples or per amplicon group and the N0 values will be based on this mean efficiency.
- Setting an individual window will only help you bring the individual efficiency closer to the mean efficiency and thus lead to a better mean efficiency. The individual N0 value depends mainly on the Ct value which is not affected by the individual window-of-linearity setting.
- Only when all samples are set to an individual window by choosing **individual fit** from the ***Window-of-Linearity*** tab (upper left) the individual efficiencies will be used in the calculation of the N0 values.
- When you have defined amplicon groups, and the ***All Samples*** graph displays all samples, you will not see the blue window-of-linearity limits because there is more than 1 window. Use the controls at the bottom left to display the amplicon groups to see the window limits per group.

Manual setting of the Nt threshold
The horizontal green line in the ***All Samples*** graph and the ***Individual Sample*** graph indicates the Nt threshold; the vertical green line indicates the Ct value. You can change the Nt threshold by drag-and-drop of the horizontal green line. The resulting change depends on the graph you are working with:

All Samples graph.
When you are in the ***All Samples*** graph, and you are using a common window-of-linearity, drag-and-drop of the green Nt threshold will change the common threshold value. However, when you have defined amplicon groups, and are displaying only one amplicon group, drag-and-drop in the ***All Samples*** graph will change the Nt threshold for that amplicon group.

Individual Sample Graph.
When you are in the ***Individual Sample*** graph and you are using a common window-of-linearity, drag-and-drop of the Nt threshold will set the common threshold. However, when you have amplicon groups defined, the drag-and-drop of the Nt threshold will set the threshold for the amplicon group in which the current sample is placed.

***NOTES*:**
- The value of the Nt threshold has no effect on the reported N0 values. The only reason to want to change the threshold value is when you want to use “comparative Ct-like” calculations with the results of LinRegPCR. In that case, you should use a common Nt threshold for all samples and/or amplicon groups

- When you have defined amplicon groups, and the ‘All Samples’ graph displays all samples, you will not see a green Nt threshold because there is more then 1 threshold. Use the controls at the bottom left to display the amplicon groups to see the Nt thresholds per group.

3.8 Handling Noisy datasets

A noisy sample is defined as a sample in which the data in the window-of-linearity are not continuously increasing. This mostly occurs when the observation noise is too large. This might be caused by the automatic exposure time setting of your qPCR apparatus. It is better to set the exposure time to a fixed value. When noisy data occur, the program can also not accurately determine the fluorescence baseline. So, you probably also have a number of samples that have a baseline error assigned to them.

The program checks for noisy data immediately after the estimation of the fluorescence baseline. When noisy data occur, the program stops the automatic processing of the data and directs you to the tab page on ‘Handling noisy data’ in the Help. A short-cut to this page is added to the ***Help*** menu. The program also shows the ***Warning*** window (Fig. 8)

moet nog naar de ppt file en voorzien worden van legend.

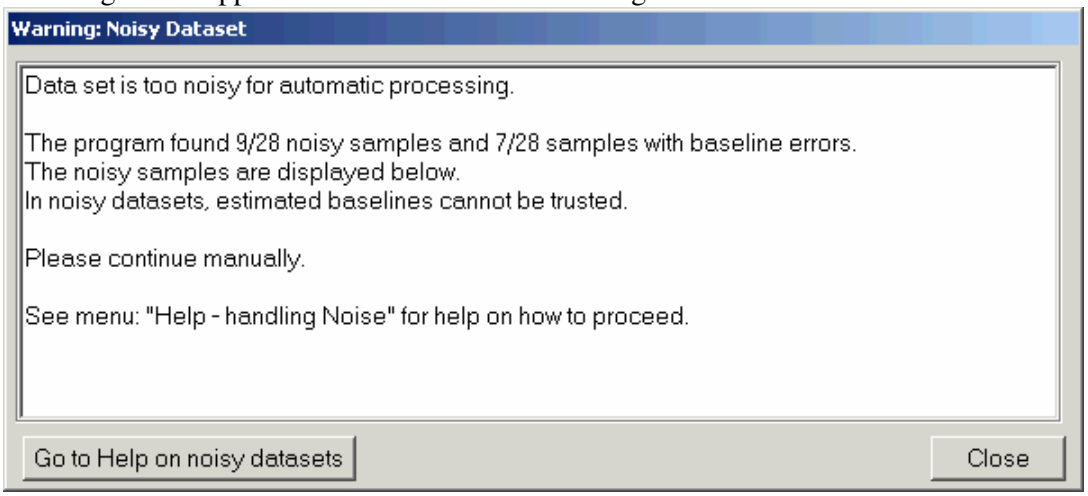

Figure 8: Warning message that pops-up when the program has detected a ***noisy sample*** during the estimation of the fluorescence baselines.

Handling Noisy data in the lab

Noisy data might occur when the distance between baseline and plateau phase is not large enough. Then the program does not have enough room to determine the baseline accurately. To solve these issues you have to go back to the lab. You can do two things to improve the data quality in the future. Firstly look at the exposure-time setting in your PCR apparatus and set this to a fixed exposure time. And secondly try to increase the baseline - plateau ratio by lowering the primer concentration in the PCR reaction.

You can try to RESCUE the current dataset by doing the analysis manually. You first have to set the correct baseline for, at least, the 'noisy' and the 'baseline error' samples. Thereafter you can set a window-of-linearity to continue the analysis. Note that the Ct values are hardly dependent on the baseline correction. So, when using the mean efficiency per amplicon and the individual Ct values, like LinRegPCR does, you may get reliable data even from a noisy

data set. But keep in mind that you are trying to rescue low quality data. Baseline errors and Noisy data stay marked in the output.

Setting a manual fluorescence baseline

To rescue this dataset you can try to set the baseline manually for the noisy samples and the samples with a baseline error. Note, however, that in a noisy dataset also samples that did pass the baseline correction may have a wrong baseline. So it is best to browse through all samples to check the baseline correction.

Use the display drop-down box on the left of the screen to select 'noisy samples'. Their PCR efficiency values are displayed at the bottom. Browsing through the samples is now restricted to those samples.

Use the manual baseline setting controls at the right of the sample graph to estimate the correct baseline (see 3.6). You have to aim for the longest set of data points on a straight line going down from the start of the plateau phase. When this line swings down, the baseline is too high, when it swings up, the baseline is too low. Only consider the points that are on a continuous line down from the plateau phase.

It may help when you first set the baseline value to the minimum fluorescence. Press '***get Min.Flu***' and then press ***apply baseline***. Press ***reset default*** when the situation became worse. The up-down arrows next to the edit field increase or decrease the baseline by 1, 0.1 or 0.01% of the current baseline. Start with steps of 1% and try to reach a straight set of data points. Fine tune this baseline with smaller steps.

You can keep an eye on the PCR efficiency; efficiencies above 2 most likely result from a baseline that is too high. Keep in mind that the PCR efficiency is dependent on the Window-of-Linearity. So check the position of the W-o-L (see below) when you start paying attention to the PCR efficiencies. When no good baseline can be set, you can exclude a sample from the analysis by checking ***Mark sample as excluded***.

NOTE:

- When you have thus set the correct baselines for the **noisy samples**, it may help to do the same for the samples that were assigned a **baseline error**. It is advisable to also check the samples that passed the automatic baseline estimation.
- When you have checked and corrected all samples you can try to set an automatic window-of-linearity by pressing the ***Recalculate W-o-L*** button.

Setting a manual Window-of-Linearity

To rescue this dataset you can try to do the analysis manually. Because of the occurrence of noisy samples', the window-of-linearity is not set at the right position. You can correct this by setting a manual window with the ***common window*** controls on the left of the screen. You have to aim for a window that gives you the least variation between PCR efficiencies (see the survey graph at bottom of the screen). When you have different amplicons in the data set you first have to define amplicon groups (***Amplicon Groups***, see 3.3) and use the controls on the ***group window*** tab to set a window-of-linearity per group.

You can press the ***Recalculate W-o-L*** button (below the manual baseline controls on the right) to set an automatic window after you have corrected the baselines in the noisy and baseline-error samples. However, in a noisy dataset this may not give the best window.

NOTE:

- The fluorescence threshold for determining Ct values is automatically set at 1 cycle below the upper limit of the W-o-L. You can choose your own threshold by entering the value in the edit field on the Nt-threshold tab (top left).

3.9 Save options

You can decide for yourself which data you save to Excel on the ***Fit / Save*** Options page that you can open by choosing ***Options - Save Options*** from the main menu. By (un)checking the boxes in front of each parameter you determine which parameters are exported. Checking the ***all columns*** checkbox will give the complete output.

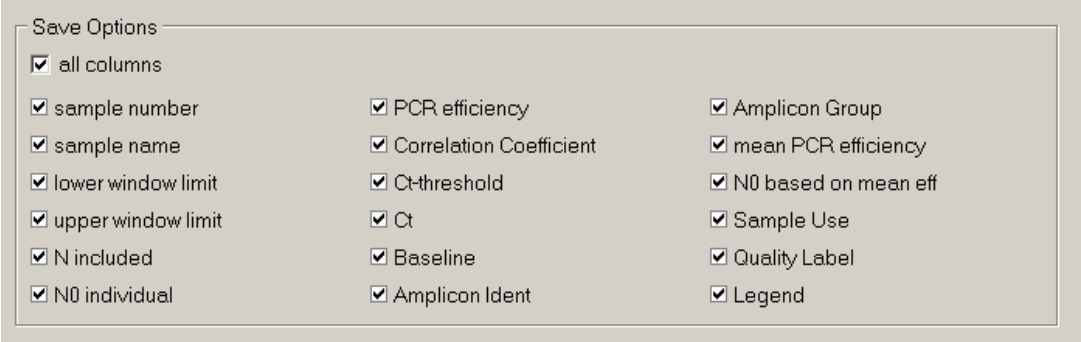

Figure 9: Save Options. By (un)checking the boxes you can decide which parameters have to be saved to Excel. Sample number, sample name, N0 based on mean efficiency, Sample use, Quality label and Legend cannot be unchecked and will always be exported.

***NOTES*:**

- The choices you make will be saved to the Windows Registry and will be used as default the next time you use LinRegPCR. This ensures that the format of the output of LinRegPCR will be the same and that you can easily apply the analysis templates you made in Excel.
- The **N0 individual** value is only included for backward compatibility. When you do use a common or amplicon window-of-linearity, you have to ignore this N0 value.

3.10 Save to Excel

Results of the quantitative PCR data analysis can only be saved to an Excel spreadsheet. To save the analysis results, choose ***File - Save to Excel*** from the main menu.
Excel has to be running!

In the ***Save to Excel*** dialog:

1. Choose to save to an existing Excel ***book*** or to a new book. By default LinRegPCR suggests to save to the book you read the data from, which is the active book.
2. Give the name of the output sheet. As default, LinRegPCR uses the name of the input sheet extended with the kind of output and the date.
3. Choose the kind of output you want to save:
   - **Complete output**: all 18 columns, also the ones that are there for backward compatibility.
   - **User defined output**: you can choose the output columns on the Fit / Save Options page. For experienced users only (see chapter 12).
   - **Compact + Complete output**: this is the recommended option when you start using LinRegPCR. The compact output page gives the columns required to further analysis of your data; the complete data is available for trouble shooting.
4. Choose the ***position*** of the output sheet(s) in the Excel book.
5. Press ***OK***

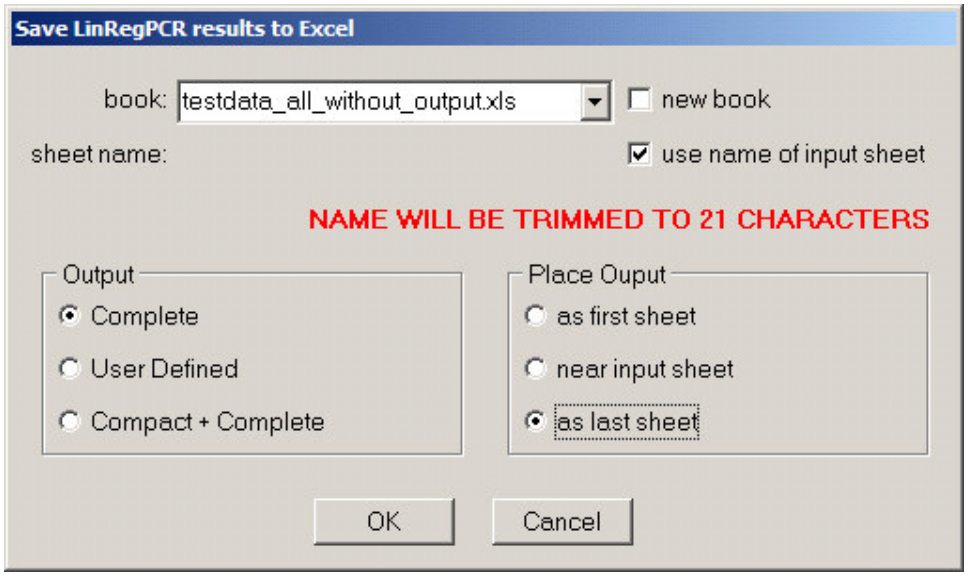

Figure 10: Save to Excel dialog.

***NOTES*:**

- Excel must be running to save results to Excel and Excel must be in 'Ready' status (see 3.1).
- When you have changed the Excel file that you read the data from during the analysis saving the results to this file may fail. In that case, saving to a new book will still be possible.
- You are not allowed to delete sheets from the active book in Excel while the ***Save-to-Excel*** dialog is open.
- When the output sheet name already exists in Excel, '_1' will be added to the name of the new file; '_2' when the '_1' is already present, etcetera.

3.11 Quality control output

The last four columns of the output look give information on the quality of the samples, as determined by LinRegPCR.

N0 = threshold /(Eff_mean^Ct)

N0	Sample_Use	Quality_checks
-9.99E+02	0 0 0	- 1 - 3 - 5 6 - - - - -
1.10E-10	0 0 3	- - - 3 - - 6 - - - - -
4.29E-10	0 0 3	- - - 3 - - 6 - - - - -
1.55E-09	0 0 3	- - - 3 - - 6 - - - - -
1.95E-09	0 0 3	- - - 3 - - 6 - - - - -
-9.99E+02	0 0 0	- 1 - 3 - 5 6 - - - - -
5.25E-09	1 2 3	0 - - - - - - - - - - -
3.58E-09	1 2 3	0 - - - - - - - - - - -
-9.99E+02	0 0 3	- - 2 - 4 - 6 - - - - -
1.58E-08	1 2 3	- - 2 - 4 - 6 - - 9 - -
2.72E-08	1 2 3	0 - - - - - - - - - - -
2.74E-10	0 0 3	- - - 3 - - 6 - - - - -
8.38E-09	1 2 3	- - - - - 5 - - - - - -
1.11E-08	1 2 3	0 - - - - - - - - - - -
1.84E-10	0 0 3	- - - 3 - - 6 - - - - -
5.46E-09	1 2 3	0 - - - - - - - - - - -
6.29E-09	1 2 3	0 - - - - - - - - - - -
-9.99E+02	0 0 0	- 1 - 3 - 5 6 - - - - -

LEGEND

Sample Use:
1: used for W-o-L setting
2: contributes to mean PCR efficiency
3: N0 value calculated
0: not used / calculated

Quality Checks:
0: passed all checks
1: no amplification
2: baseline error
3: no plateau
4: noisy sample
5: PCR efficiency outside 5%
6: excluded from mean Eff
7: excluded by user
8: included by user
9: manual baseline

if amplicon groups are defined
then rules are applied per group

Figure 11: Example of the last 4 columns of the output to Excel showing the quality check output

The ***Sample_Use*** column tells you whether the sample is used for (1) setting the window-of-linearity, (2) calculating the mean efficiency, and (3) whether a starting concentration (N0) is calculated. When either of these is not applicable the column contains a '0'. When no starting concentration is calculated, the **N0** or **N0_(mean eff)** column will contain '-999'. If you want to calculate the N0 by hand you have to use the equation above the **N0** column. All required data are present in the compact and the complete output. Make sure you have the **Nt-threshold**, **Ct-value** and **mean PCR efficiency** columns in your user-defined output (see 3.9) to be able to calculate N0 manually.

The ***Quality_checks*** column contains 0 when the sample is OK. The other 9 positions can contain values of 1 thru 9 of which the explanation is given in the ***LEGEND*** column under **Quality Checks**. Based on these values you can decide to include or exclude the N0 value from further analysis. Samples without amplification (**1**), usually also do not reach the plateau (**3**) and are excluded by the program (**6**). Similarly, samples that are assigned **baseline error** (**2**), usually have a deviating PCR efficiency (**5**) and are always excluded by the program (**6**), unless you decide to set a manual baseline (**9**) which results in a N0 value being calculated. Samples that do not reach the plateau phase (**3**) OR have a deviating PCR efficiency (**5**) are not included in the calculation of the mean efficiency but a N0 value is calculated. However, when both conditions apply (**3 5**) the samples are rejected by the program (**6**) and receive a N0 of -999. When you do not agree with this decision you can overrule this by calculating the N0 value in Excel (see above)

4. Interface components

4.1 Menu and Submenus

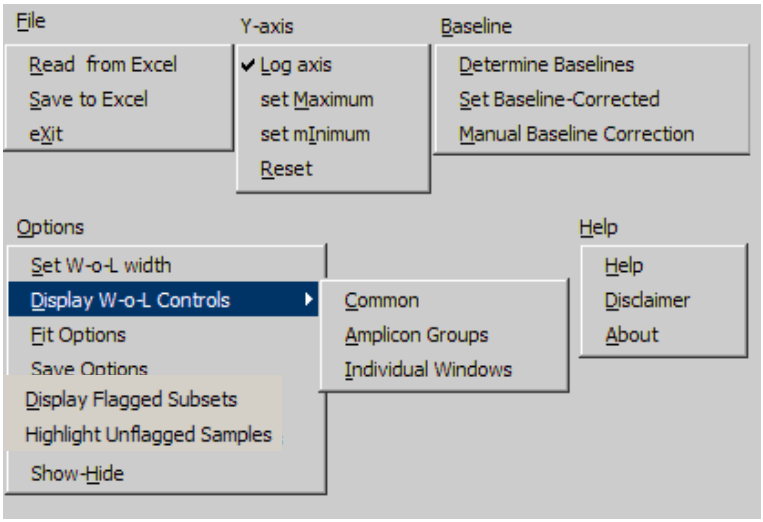

Figure 12: Main menu and sub-menu choices.

File Menu

Read from Excel:

Use this option to read data from Excel (see 3.1)

Save to Excel:

Use this option to save results to Excel (see 3.10)

Exit:

Closes the program without saving any results or choices.

Y-axis menu

Log axis:

You can use this option to toggle between a logarithmic Y-axis (default and recommended) and a normal Y-axis. The check boxes on the graphs have the same effect.

Set Maximum / Set Minimum:

This choice will open a dialog window in which you can enter the Y-axis maximum or minimum. Depending on the status of **Log axis** (see above) the input has to be a log-value or a normal value.

Reset:

This choice resets the Y-axis to the default values that were determined automatically after the **Read from Excel** and again after the baseline correction.

Baseline Menu

Determine baselines:

Has the same effect as pressing the ***Determine Baseline*** button (see 3.2).

Set Baseline Corrected:

This option enables you to skip the baseline subtraction procedure when you had forgotten to indicate in the Read-from-Excel dialog that the PCR apparatus had already done a baseline subtraction.

Manual Baseline:

This option opens the manual baseline controls that allow you to manually set or correct the baseline (see 3.6 and 3.8).

Options menu

Set W-o-L width:

By default 4 data points are included in the window-of-linearity. You can also change this number to your preference on the Fit / Save Options Tab page, but using fewer points is not recommended because it increases the variation in PCR efficiency values.

Display W-o-L Controls:

In the submenu of this option you can choose to open the controls that you need to manually set the upper and lower limit of the window-of-linearity. You can choose to set the common, amplicon group and individual windows (see 3.7 and 3.8).

Fit Options:

Choosing this option opens the Fit / Save Options page. On this page you can set the fit options for the individual window-of-linearity. Note that using the individual windows is not recommended. This option is maintained for backward compatibility.

Save Options:

Choosing this option opens the Fit / Save Options page. On this page you can choose the output columns for the User-Defined output. These choices will be saved to the Windows Registry and be automatically used the next time you use the program (see 3.9).

Display Flagged Subsets:

This option opens a list box (on the middle left of the screen) from which you can choose to display the samples that the program has flagged as: **no-amplification**, **baseline error**, **no-plateau, noisy sample,** or **efficiency outlier**. The first three categories are assigned by the baseline estimation procedure, the "efficiency outliers" are determined while saving the results to Excel (see 3.11).

Highlight Unflagged Samples:

Choosing this option displays the unflagged samples (see above) in green and the other samples in red in the PCR efficiencies graph at the bottom of the screen.

Show-Hide:

Opens the Show-Hide page on which you can manually choose which samples to display in the "All Sample" graph. This choice does not affect the calculations or the output and is overruled when you choose to display separate amplicon groups.

Help menu

Help:

Opens the **Help** window.

Help on Noisy data sets:

This menu item is added when a noisy sample is encountered (see 3.8)

Disclaimer:

Opens the **Disclaimer** window. This window is also displayed at first use of the program.

About:

Opens the **About** box with the version information and a link to the LinRegPCR web page.

4.2 Tab pages

All Samples

This page shows the amplification curves of all samples. The blue lines indicate the upper and lower limit of the W-o-L; the green line is the Nt threshold. Points in the W-o-L are shown in red. When you have defined amplicon groups, which have different W-o-L and Nt you do not see those lines. The data points in the W-o-L of the sample shown at the ***Individual Sample*** page are shown in green.

Individual Sample

This page shows the amplification curve of the current sample. Lines and colors are as in the ***All Samples*** graph. When the sample is fitted to a common W-o-L the data points in the W-o-L are shown in red, when the W-o-L is set per amplicon group the data points are orange. Yellow data points indicate an individual W-o-L.

Amplicon Groups

On this page you can define the amplicon groups and set the criteria for exclusion of samples from the calculation of the mean PCR efficiency (see 3.3 and 3.5).

Data

This page contains 3 sub-pages:

input data

This Tab page shows the imported data. The only function of this page is to check that all data have been imported.

valid data

This Tab page shows the baseline corrected data with the negative (below baseline) data removed. You can save these data to Excel. On a later date you can import them again into LinRegPCR without going though the baseline correction again. This is especially handy when you had to do a lot of manual baseline setting. Import them as ABI or Lightcycler480 format and check **Yes** at baseline corrected on the ***Read from Excel*** dialog.

sample data

This page shows the data of the current sample, which is the sample displayed in the 'Sample Data' tab. The data points labeled 'Y' are the ones included in the window-of-linearity and used for the calculation of the individual PCR efficiency.

Save Options

The top of this page determines the rules for fitting individual samples. These are no longer used when you set a common W-o-L or a W-o-L per amplicon group. When you choose to fit individual W-o-L's these fit options are used. This part of the page is just there for backward compatibility.

The bottom part of this page contains the checkboxes for selecting of the columns you want to see in the output of the analysis results to Excel (see 3.9).

Show or hide

With the list boxes on this page you can select which samples are shown in the graph on the **All Samples** page and in the **Survey graphs** (see 6). This can be handy when you want to show amplification curves of a selection of your data in a presentation. The choices made on this page are overruled when you browse through amplicon groups and when you select marked subsets for display.

4.3 Survey graphs

At the bottom of the LinRegPCR screen there is a notebook containing three Tab pages displaying a survey of all PCR efficiencies, correlation coefficients and number of included points, respectively. In each of these graphs you can click on a sample to highlight this sample in the ***All Samples*** graph. It will be shown as green points. Alternatively, you can inspect the chosen sample in the ***Individual Sample*** graph. This is the easiest way to look for deviating samples.

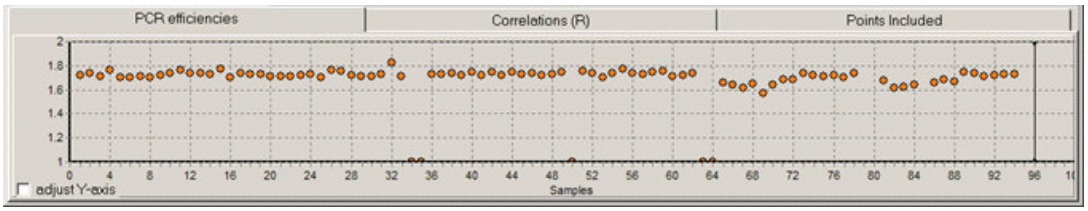

Figure 13: Notebook with Tab pages to display the survey graphs of PCR efficiencies, correlation coefficients and number of included points.

The data points in the survey graphs are red when you are using a common window, orange when you have defined amplicon groups and yellow when a sample is fitted to an individual window.

Exporting graphs

When you click the right mouse button on a graph, the graph is copied to the Windows Clipboard as a bitmap. This bitmap can be pasted into your presentation.

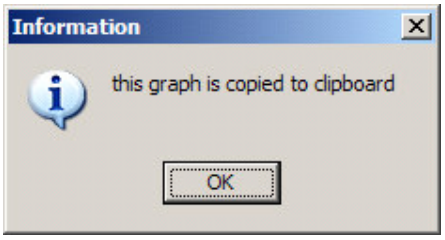

Figure 14: Notification that the graph has been copied to the Windows Clipboard.

5. Frequently asked questions

Please use the button on the About box and in the Help of the program to access the LinRegPCR website and look at the FAQ's (http://LinRegPCR.nl).

6. Acknowledgements

We wish to thank Prof. Antoon Moorman (Heart Failure Research Center, AMC, Amsterdam) and Dr. Onno Bakker (Department Endocrinology and Metabolism, AMC, Amsterdam) for their advice and support. All qPCR users in the Heart Failure Research Center and the department of Neurogenetics (AMC, Amsterdam) are acknowledged for their critical comments and suggestions when using beta-versions of this program.