Quantitative data generation for systems biology: the impact of randomisation, calibrators and normalisers

M. Schilling, T. Maiwald, S. Bohl, M. Kollmann, C. Kreutz, J. Timmer and U. Klingmüller

Abstract: Systems biology is an approach to the analysis and prediction of the dynamic behaviour of biological networks through mathematical modelling based on experimental data. The current lack of reliable quantitative data, especially in the field of signal transduction, means that new methodologies in data acquisition and processing are needed. Here, we present methods to advance the established techniques of immunoprecipitation and immunoblotting to more accurate and quantitative procedures. We propose randomisation of sample loading to disrupt lane correlations and the use of normalisers and calibrators for data correction. To predict the impact of each method on improving the data quality we used simulations. These studies showed that randomisation reduces the standard deviation of a smoothed signal by 55% $\pm$ 10%, independently from most experimental settings. Normalisation with appropriate endogenous or external proteins further reduces the deviation from the true values. As the improvement strongly depends on the quality of the normaliser measurement, a criteria-based normalisation procedure was developed. Our approach was experimentally verified by application of the proposed methods to time course data obtained by the immunoblotting technique. This analysis showed that the procedure is robust and can significantly improve the quality of experimental data.

1 Introduction

Blotting techniques are widely used to analyse components in biological systems. They are based on the separation of components according to the molecular weight within a gel and transfer to a membrane followed by a detection process. The presence of proteins and/or their modifications in complex mixtures is examined by immunoblotting using specific antibodies in combination with chemiluminescence detection. So far, the data generated by immunoblotting have been primarily qualitative but the recent report of data-based mathematical modelling of the JAK-STAT signalling pathway [1] demonstrates the potential of using quantitative immunoblotting for systems biology approaches.

Here, we suggest new methodologies to improve data acquisition and data processing for quantitative immunoblotting. We propose randomised gel loading to transform correlated blotting errors into uncorrelated blotting errors by loading samples on the gel in a non-chronological order: Neighbouring lanes on the gel are used, not for consecutive time points, but in a randomised way. Furthermore, we suggest normalisation using data of calibrators (purified proteins of a different molecular weight from that of the protein of interest and the same antibody binding epitope added to cell lysates prior to immunoprecipitation) and normalisers (endogenous proteins quantified by reprobing the immunoblot). Smoothing splines of the normaliser or calibrator signals were employed to correct immunoblotting data in an unbiased, criteria-mediated framework.

In addition, we present a quantitative analysis of the effects of these improvements by assessing the influence of each method on the standard deviation and correlation structure of simulated data and measurement processes. By applying our procedures to a data set comprising five independent experiments, each measuring eight protein species with 20 time points, we validated the performance of our approach method under experimental conditions.

2 Randomisation reduces standard deviation of immunoblotting data more than two-fold

Simulations of typical immunoblotting experiments were performed by generating a simulated signal x^* with quadratic rise and exponential decay and a maximum at half gel slot number, equidistantly sampled (Fig. 1*b*). This simulates a typical time course experiment after stimulation with a hormone. The true signal x^* was processed as follows:

First, a multiplicative, uncorrelated pipetting error of strength σ was applied, as shown in Fig. 1*a*, representing errors derived from unequal cell number or errors in pipetting the cellular lysates:

$$x' = x^* \cdot (1 + \varepsilon) \quad \varepsilon \sim N(0, \sigma)$$

Secondly, a multiplicative, strongly correlated blotting error was applied, representing errors from differences in migration in the SDS polyacrylamide gel or unequal transfer to the membrane, a common and probably underestimated problem in immunoblotting (compare the estimated blotting error in Figs. 4*a* and 6*a* and *d*)

$$x = x' \cdot g$$

IEE Proceedings online no. 20050044

doi:10.1049/ip-syb:20050044

Paper first received 1st July and in revised form 7th September 2005

M. Schilling, S. Bohl and U. Klingmüller are with the German Cancer Research Center, Im Neuenheimer Feld 280, Heidelberg D-69120, Germany

T. Maiwald, M. Kollman, C. Kreutz and J. Timmer are with the Freiburg Centre for Data Analysis & Modelling, University of Freiburg, Eckerstr. 1, Freiburg D-79104, Germany

M. Schilling and T. Maiwald contributed equally to this work.

E-mail: m.schilling@dkfz.de

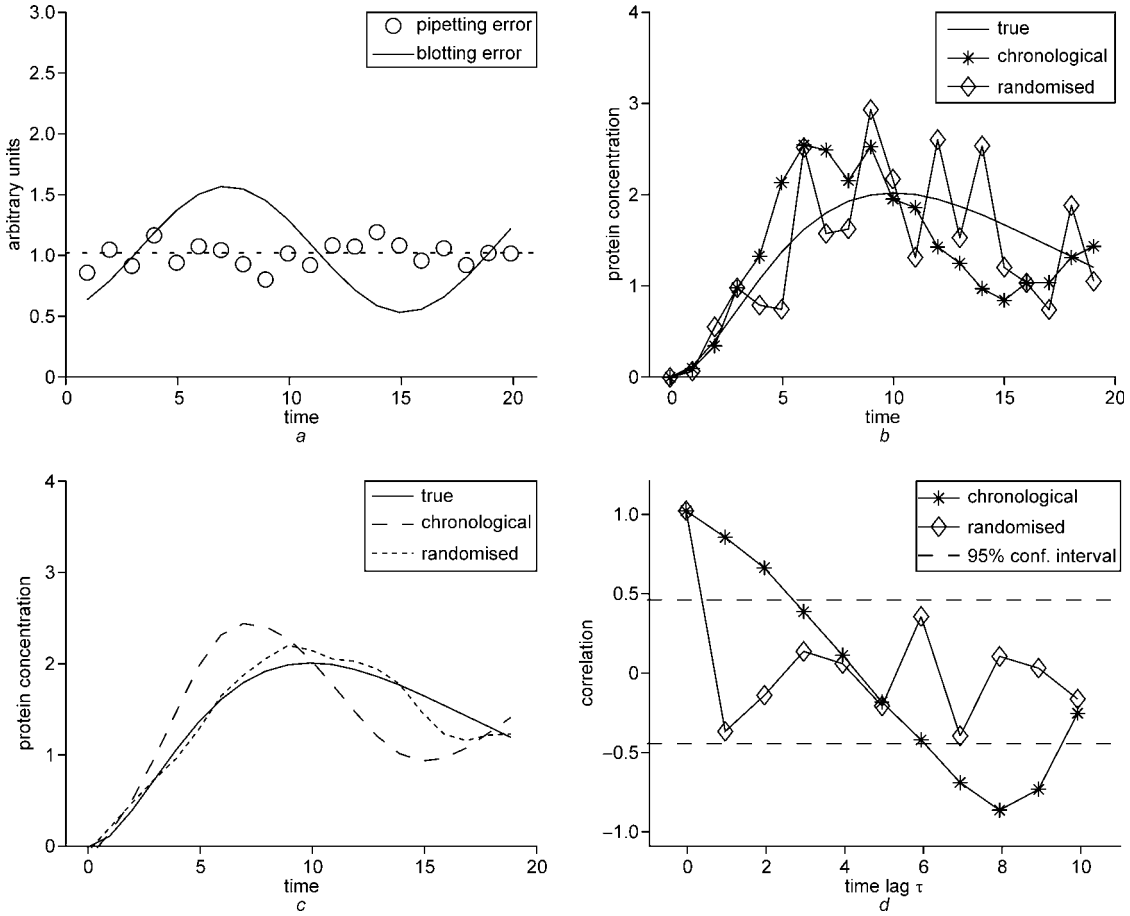

Fig. 1 *Effect of randomisation on immunoblotting data*

a Simulated uncorrelated pipetting error and highly correlated, sine-like blotting error
b Simulated signal perturbed with pipetting and blotting error in chronological and randomised manner
c Only randomised procedure does not change characteristics of true signal, as smoothed data show
d Residuals of perturbed to true signal exhibit strong autocorrelation for chronological procedure, which is not agreeable with white noise. Randomisation prevents this autocorrelation

with the blotting error g represented by a sine function with mean 1 and varied phase, amplitude and frequency.

The processing was applied to a chronological signal and to a randomised true signal, respectively, leading to simulated measurements such as in Fig. 1*b*. Note that the chronological signal is rather smooth but changes the characteristic of the true signal: the maximum occurs earlier, and a new minimum is observed at $t = 15$. The randomised signal, on the other hand, is very noisy, but does not introduce misleading effects. The smoothed, processed, randomised signal is very close to the true one, whereas the smoothed, processed, chronological signal conserves the correlated deviations from the true signal (Fig. 1*c*).

The correlation structure of the deviations can be investigated through the autocorrelation function (Fig. 1*d*). For uncorrelated errors, the autocorrelation function should drop from 1 at $\tau = 0$ into the 95% confidence interval for $\tau > 0$ [2]. This is not the case for the processed chronological signal, which can lead to incorrect conclusions if methods assuming uncorrelated noise are applied. Besides visual inspection of the autocorrelation function, the improvement of data quality can be quantified by the error reduction factor by randomisation, defined as the reduction of the standard deviation of the smoothed signal by means of randomisation

$$\textit{error reduction factor by randomisation} = \frac{\textit{standard deviation (smoothed chronological data)}}{\textit{standard deviation (smoothed randomised data)}}$$

For the illustrated data set, an error reduction factor of 0.4 was calculated, i.e. the standard deviation could be decreased by 60%. The reduction can only be quantified when the true data are available, which is certainly not the case in real measurements. Hence, the question arises whether a general reduction factor can be established by randomising or whether it depends on experimental parameters such as the number of lanes, maximum signal strength, blotting error or pipetting error. We performed a simulation study showing that, for small pipetting errors, a reduction to 45% $\pm$ 10% could be established independently from other parameters.

3 Simulation study

Several parameters were varied quantitatively to assess the effect of randomisation. This included the number of lanes (10–100), the number of sine periods of the blotting error (0.8–2.2), the strength of the blotting error (ratio of smallest to largest value ranging from 1.5 to 10), the strength of the pipetting error (σ ranging from 0 to 1) and the maximum signal strength (0.1–20).

During the variation of one parameter, the other parameters were fixed:

- number of lanes: 20
- number of sine periods of the blotting error: 1
- strength of the blotting error (max/min): 3
- standard deviation of the pipetting error: 0.1
- maximum signal strength: 2.

Figure 2 displays the error reduction factor for all parameter variations. Variation of lane number (Fig. 2*a*) showed that randomisation is recommended for 15 or more lanes. For the other investigated parameter ranges, no strong effect could be observed for all variations (Figs. 2*b*–*d*), except for the strength of the pipetting error (Fig. 2*e*).

As pipetting errors are uncorrelated, they cannot be reduced by randomisation: if the fraction of the pipetting errors increases, the randomisation has less effect. In general, randomisation decreases the standard deviation in quantitative immunoblotting by 55% ± 10% of the value without randomisation, as long as the standard deviation of the pipetting error does not exceed 20%. An approach

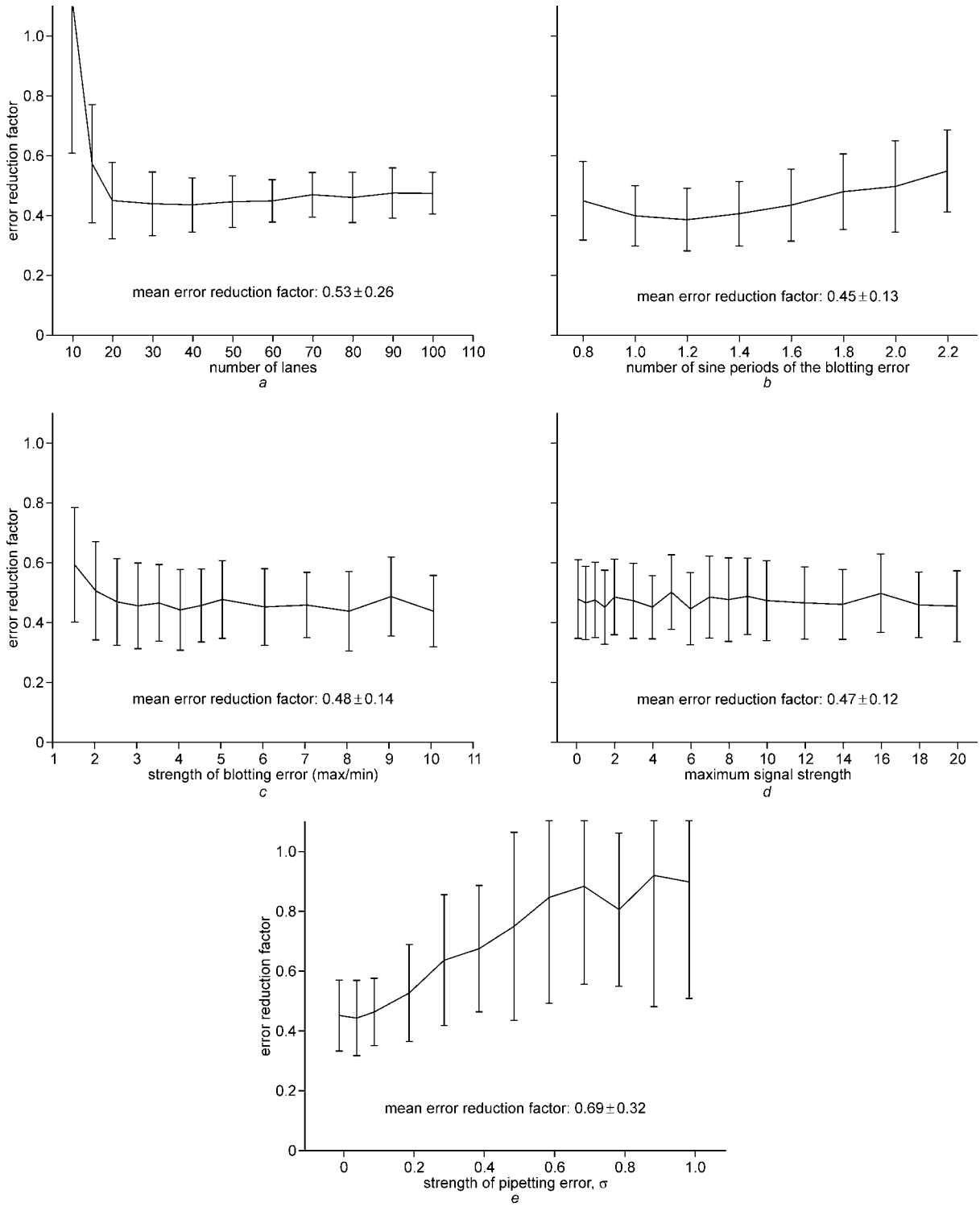

Fig. 2 *For gels with more than 15 gel lanes, randomisation reduces standard deviation robustly by 50%*

a Variation of lane number
b Variation of blotting error characteristics
c Variation of error strength
d Variation of signal strength
e Variation of pipetting error
Increasing pipetting error to blotting error ratio decreases error reduction factor, as only blotting errors are tackled by randomisation

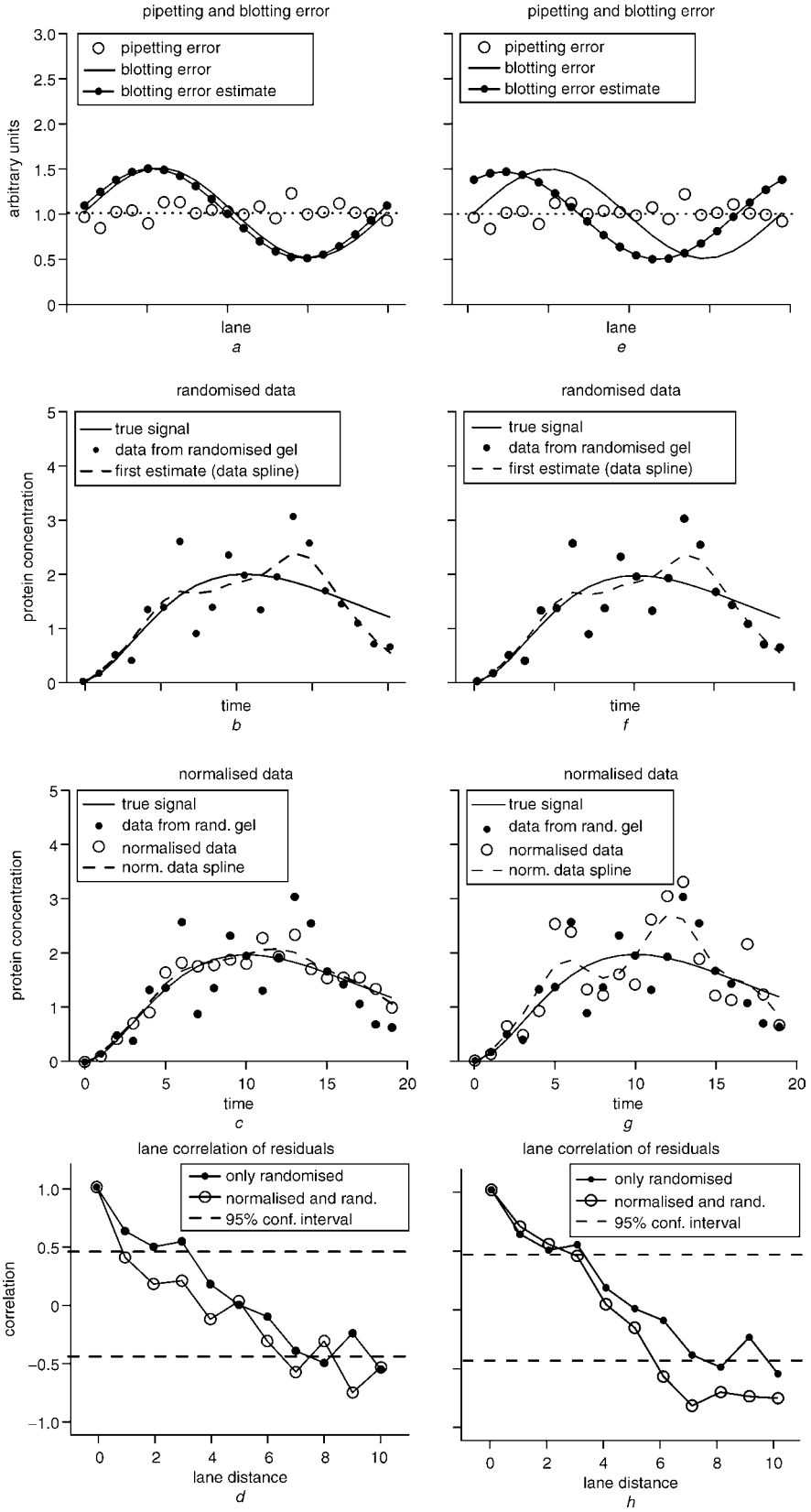

Fig. 3 *Normalisation of simulated time course data*

Left panel: valid procedure according to our criteria
a Blotting error is well estimated corresponding to suitable normalisation protein
b Perturbation of simulated signal
c Perturbation is strongly reduced after normalisation
d Correlation in gel domain of residuals is improved
Autocorrelation, i.e. correlation in time domain, agrees for both randomised signals with white noise (not shown)

Right panel: procedure rejected according to our criteria
e Blotting error estimate is phase shifted, corresponding to too distant normalisation protein
f Perturbation of simulated signal
g Perturbation cannot be reduced with normalisation
h Lane-to-lane correlation is not improved

to control the pipetting error in experiments is sampling the same number of cells for each time point or measuring and adjusting total protein concentration.

4 Criteria for employing calibrators and normalisers further to improve quantitative immunoblotting data

Calibrators and normalisers possess a constant concentration. Fluctuations occur only as measurement errors. As the blotting error changes gradually from lane to lane, and other errors such as the pipetting error are rather uncorrelated, the blotting error can be estimated by smoothing of the calibrator or normaliser signal, e.g. with a smoothing spline [3]. The smoothing is carried out with a cubic spline approximation, the smoothness being determined by generalised cross-validation. Based on this blotting error estimate, the protein of interest can be normalised by division. However, as the blotting error is a local property of the gel, normalisers and calibrators with a similar molecular weight to that of the protein of interest are required. If the position of the normaliser is too distant on the blot, or its signal is too noisy, the normalisation procedure can even be detrimental to the data. We therefore developed criteria for employing normalisers and calibrators. The following discussion applies equally well to normalisers and calibrators.

Figure 3*a* shows a simulated blotting error and a good estimation, corresponding to the smoothed signal of an appropriate normaliser in a real experiment. Smoothing

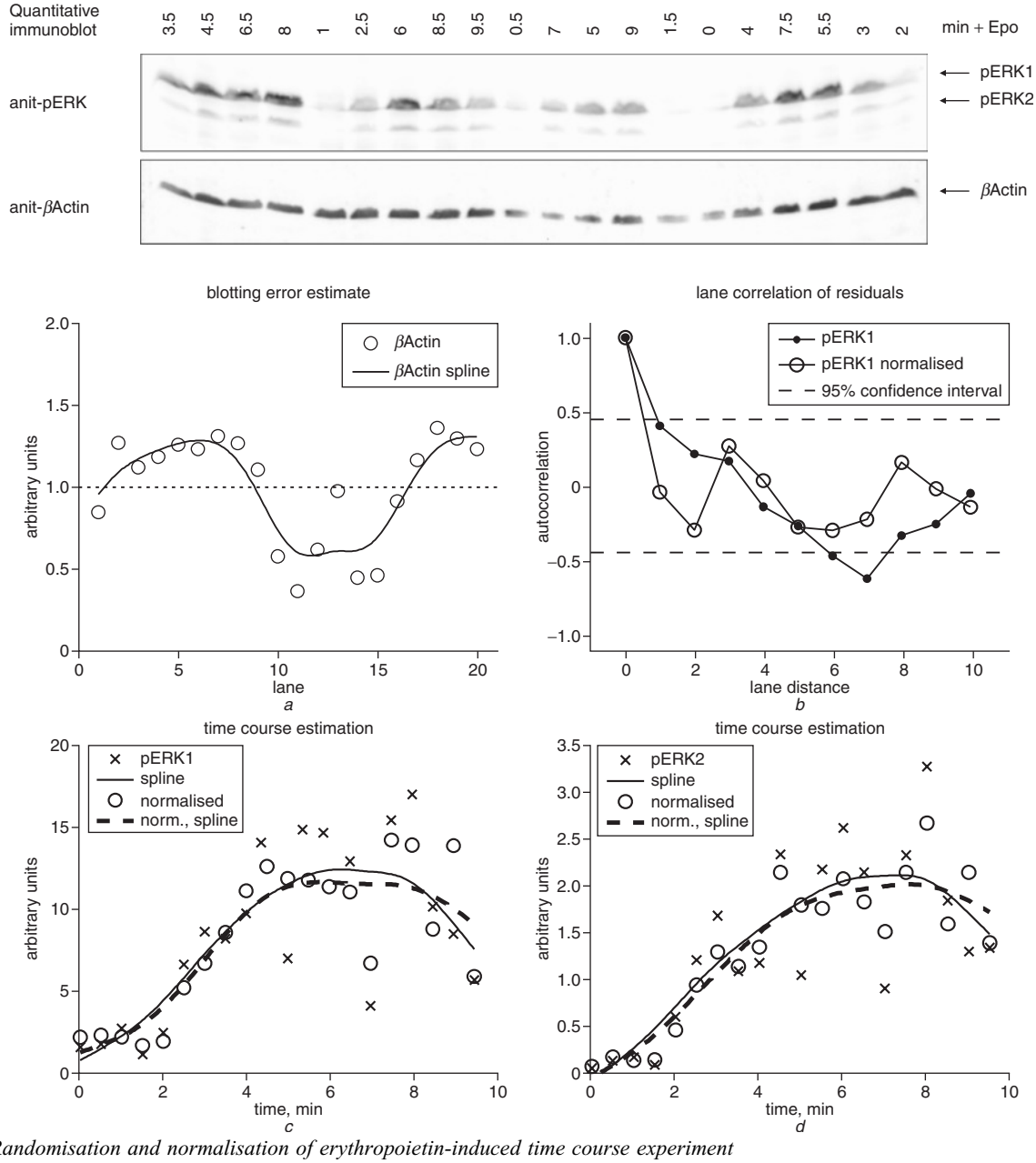

Fig. 4 *Randomisation and normalisation of erythropoietin-induced time course experiment*

BaF3-EpoR cells are stimulated with 50 units ml^{-1} Epo resulting in ERK phosphorylation. Gel electrophoresis has been performed with randomised, non-chronological gel loading with βActin as normaliser protein (upper panel)

a Smoothed measurements of βActin serve as estimate of strong, sine-like blotting error

b Normalisation destroys autocorrelation

Normalisation significantly reduces standard deviation of pERK1/2 measurements compared with spline-smoothed pERK1/2 signal (solid line), which serves as first estimate of true signal

c pERK1

d pERK2

the processed randomised signal leads to an acceptable estimation of the true signal (Fig. 3*b*). Smoothing the normalised signal yields virtually the true signal itself, as shown in Fig. 3*c*. Even the correlation structure of the estimation error in the gel domain is improved (Fig. 3*d*). The estimation of the blotting error displayed in Fig. 3*e* is inaccurate: a strong phase shift can be observed, corresponding to a normaliser measurement of a different molecular weight from the protein of interest. The blotting error can be simulated by a sine wave surface plot covering the entire blot area, resulting in a phase shift of the blotting error estimate when moving vertically away from the position of the protein of interest. In this situation, normalising the data increases the deviation of the estimated signal from the true one (Fig. 3*g*). Hence, a criterion for whether normalisation is applicable would be a decreased standard deviation of the estimated signal. Unfortunately, this requires knowledge of the true curve, which is not given. Instead, the smoothing spline curve of the randomised but not yet normalised signal is used as first estimate of the true signal. If the normalisation procedure is valid, a new spline curve can be calculated based on the randomised and normalised data; otherwise, the former values are kept. The error reduction factor by normalisation is defined as

$$\textit{error reduction factor by normalisation} = \frac{\textit{standard deviation of data from first estimate}}{\textit{standard deviation of normalised data from first estimate}}$$

The shown simulated data sets posses the following standard deviations:

(*a*) Case 1 (Figs. 3*a*–*d*):

- standard deviation of the data from the true signal: 0.533
- standard deviation of the data from first estimate: 0.722
- standard deviation of the normalised data from the true signal: 0.157
- standard deviation of the normalised data from first estimate: 0.515
- error reduction factor: $0.515/0.722 = 0.7133$

(*b*) Case 2 (Figs. 3*e*–*h*):

- standard deviation of the data from the true signal: 0.533
- standard deviation of the data from first estimate: 0.722
- standard deviation of the normalised data from the true signal: 1.208
- standard deviation of the normalised data from first estimate: 1.068
- error reduction factor: $1.068/0.722 = 1.4792$

The estimated standard deviation improves in case 1 (error reduction factor <1) but becomes worse in case 2 (error reduction factor >1). Hence, the normalisation procedure is only applicable in the first case, reducing the true standard deviation from 0.533 to 0.157, i.e. to 30%. This procedure works robustly for normalisers and calibrators, as long as the immunoblot is randomised.

5 Application to experimental time course

The randomisation and normalisation procedure was applied on an erythropoietin-induced time course experiment resulting in phosphorylation of ERK1 and ERK2. Samples were loaded in a randomised order and separated on a 17.5% SDS polyacrylamide gel. Membranes were developed with chemiluminescent substrates and quantified with a CCD camera (Fig. 4). We calculated the standard deviation of the signals from their spline approximation as 2.524 for pERK1 and 0.455 for pERK2. Normalisation with βActin reduced the standard deviation to the spline approximation to 1.878 (74%) for pERK1 and 0.262 (58%) for pERK2. The reduced lane correlation for the normalised data confirms the quality of data processing. In this case, the correlation structure of the systematic blotting error could be disrupted, thus validating the normalisation.

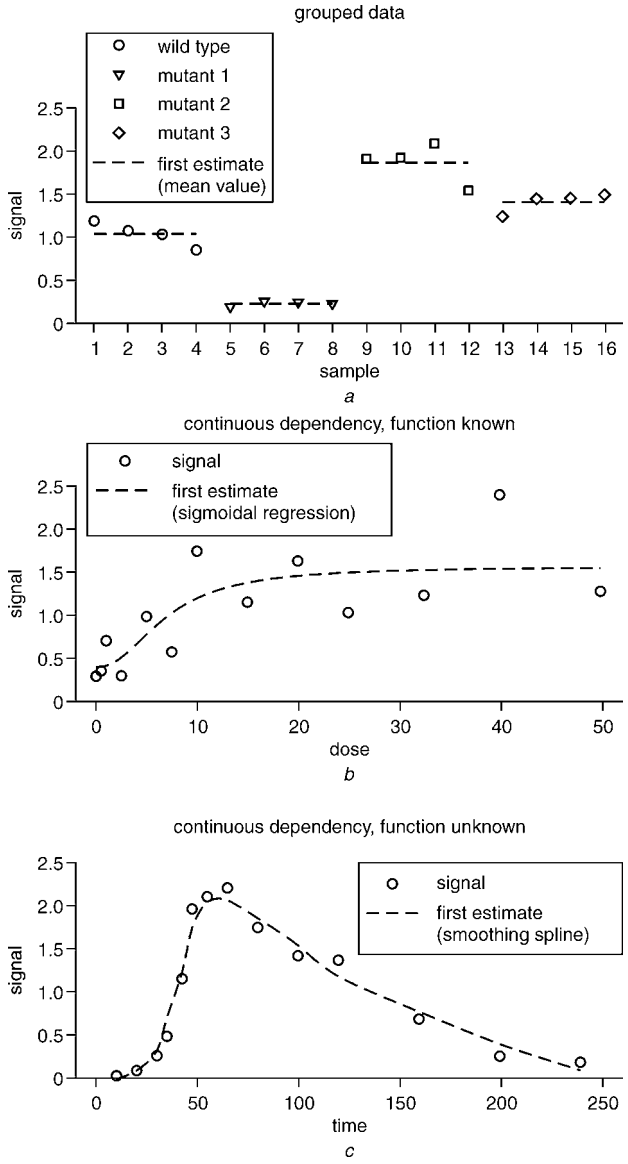

Fig. 5 *First estimates used for criteria-mediated error reduction*

a In grouped data, such as mutant to wild type comparisons, first estimate is calculated as mean value of samples loaded in replicates
b If known continuous dependency between data points exists, first estimate is calculated as regression function; for example, sigmoidal regression estimates dose-response experiment
c In cases where function is unknown, including time course experiments, first estimate consists of smoothing spline
These are artificial data for illustration purposes

6 First estimates used for criteria-mediated error reduction

The proposed criterion for error reduction needs a first estimate, which is compared with the measured and the normalised data, respectively. Above, we discussed time course data, where a smoothing spline adequately describes

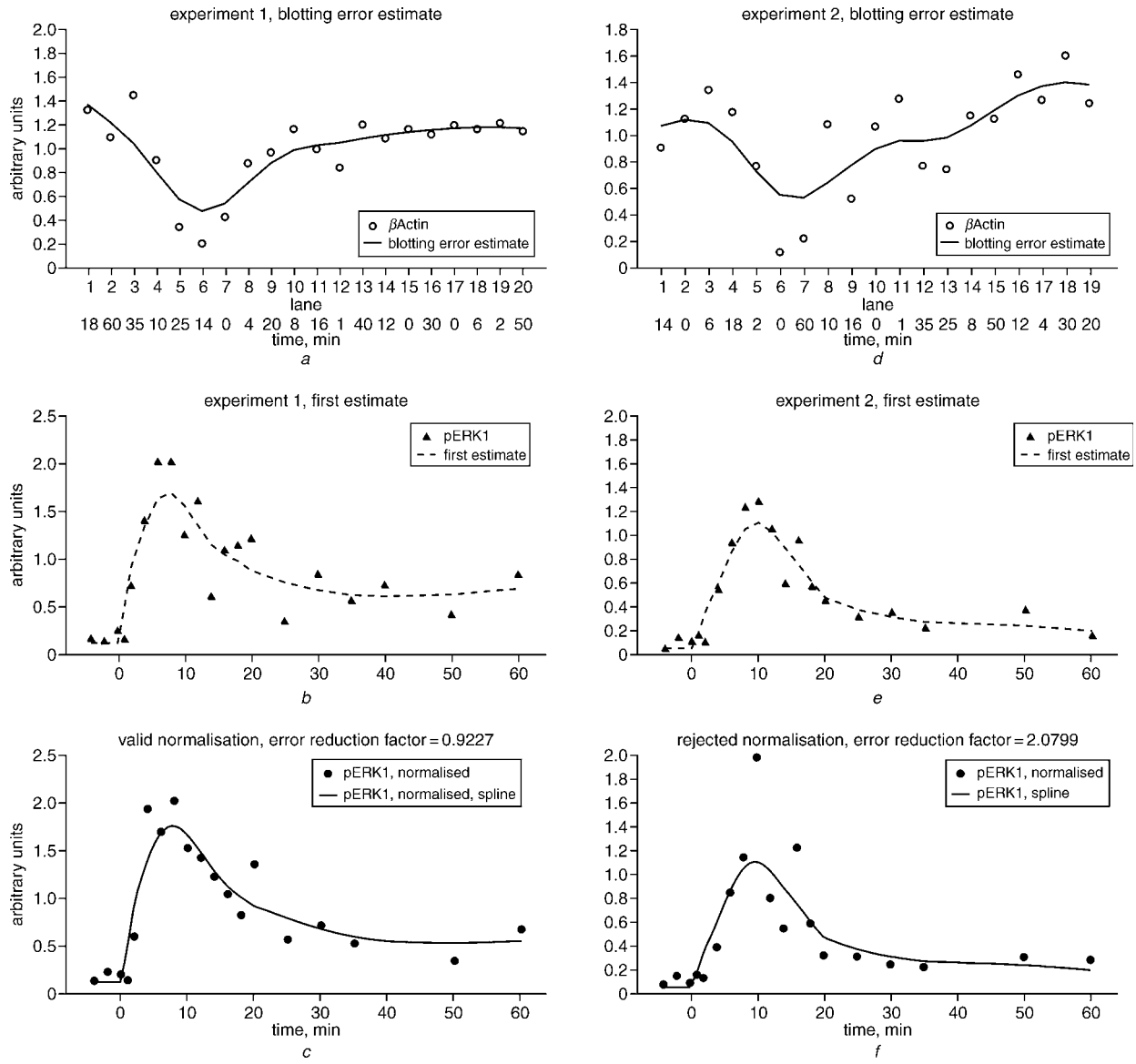

Fig. 6 *Normalisation of experimental quantitative immunoblotting data*

Left panel: valid procedure according to our criteria
a Data points of βActin and smoothing spline serving as blotting error estimate
b Smoothing spline (dashed line) serves as first estimate for data points of pERK1, which are normalised using blotting error estimate shown in *a*
c As normalised values are closer to first estimate (error reduction factor = 0.9227), normalised values are kept, and new smoothing spline is calculated (solid line)

Right panel: procedure rejected according to our criteria
d Data points of βActin and smoothing spline serving as blotting error estimate
e Smoothing spline (solid line) serves as first estimate for data points of pERK2, which are normalised using blotting error estimate shown in *d*
f As normalised values are not closer to first estimate (error reduction factor = 2.0799), normalisation is rejected, and original data are retained

the unknown functional dependency between samples. However, our procedures can also be applied to other experimental settings. Therefore we developed three categories for first estimates, as shown in Fig. 5

(i) In grouped data, as in mutant-to-wild type comparisons, the first estimate is the mean value of replicates.
(ii) In experiments with a known continuous functional dependency between time points and known function, such as dose response assays, the first estimate is calculated by a regression function.
(iii) For experiments with an unknown continuous functional dependency between time points, including time course analysis, the first estimate is represented by a smoothing spline.

Using this approach, we are able to process data derived from any immunoblotting experiment robustly. Care has to be taken if a regression is used as a first estimate. It is important to show in advance that the signal behaves as expected to prevent incorrect use of the criterion. If there is uncertainty, a smoothing spline might be more appropriate, the only prerequisite being a smooth signal behaviour.

7 Data processing of time course experiments in murine cell lines

To analyse the robustness of our data processing methods, we applied the standard operating procedures and data normalisation criteria to a large, yet noisy, data set obtained by quantitative immunoblotting of mouse cellular lysates. The murine BaF3 cell line was transfected with five different HA-tagged EpoR variants and stimulated for 1 h with 50 units ml^{-1} Epo. A total of 20 samples were taken at regular intervals so that the changes in phosphorylation and

total amount of four different proteins (the Epo receptor (EpoR), Janus kinase 2 (JAK2) and extracellular regulated kinase 1 (ERK1) and 2 (ERK2)) could be followed. EpoR and JAK2 were enriched by immunoprecipitation, and ERK1 and ERK2 were analysed from total cellular lysates. SDS gel electrophoresis was performed with randomised sample loading, and phosphorylation levels were measured using quantitative immunoblotting with phospho-specific antibodies. Furthermore, total amounts of these proteins were determined by reprobing of the membranes with the respective antibodies. Calibrators of EpoR and JAK2 added to the lysates prior to immunoprecipitation were quantified in parallel. The membranes used for measuring ERK1 and ERK2 were reprobed once more, and βActin was measured, serving as normaliser for these proteins.

After quantification of all proteins with the LumiImager system, we applied criteria-mediated normalisation procedures to the data. As the data generated were very noisy, we tested whether our methods could improve their quality. We calculated the error reduction factor of the standard deviation of the values compared with a smoothing spline, which served as first estimate. In 33% of all cases, the error reduction factor was smaller than 1, resulting in a valid normalisation procedure. In all other instances, the error reduction factor was larger than 1. In these cases, normalisation was rejected and the original data were retained.

Figure 6 shows two examples of criteria-mediated error reduction. In the left panel, the normalisation was valid, with the corrected values closely following the smoothing spline (Fig. 6*c*). In the right panel, normalisation was rejected, and the original data were kept. The most common reason for rejected normalisation was poor quality of the normaliser or calibrator measurement, as evidenced by the deviation of data points of the normaliser βActin from the blotting error estimate (Fig. 6*d*). However, as immunoblotting was performed by randomised gel loading, no misleading effects occurred that could lead to false interpretation. For visual investigation, the numbers were approximated with a smoothing spline, displaying convincing time course dynamics (Fig. 6*f*). By comparing the data obtained with previous experiments, we could show that our data processing methods substantially improved the reliability of measurements in an unbiased manner.

8 Conclusions

We show that, by the suggested data processing procedures, the standard deviations of data generated by quantitative immunoblotting can be decreased to approximately 55%, thereby increasing data quality substantially. This is largely independent of the experimental setting and quality of the immunoblotting procedure. In contrast to traditional chronological sample loading, we demonstrate the benefit of randomised immunoblots. Randomisation is most useful if the signal and blotting error are in the same frequency range. As the number of measurements is often highly limited in biochemical experiments, the sampling of time courses is rather coarse. This leads to similar frequencies of the signal and the blotting error. By quantification and plotting of the signal intensity against time, the same information is obtained as with chronological immunoblots; however, as the correlation between neighbouring lanes is disrupted, the standard deviation of the smoothed signal is reduced more than two-fold, leading to data of higher quality and therefore to fewer experiments being necessary for novel biological insight.

Furthermore we demonstrate that calibrators and normalisers allow for correction of immunoblot data, provided there is a good estimate of the blotting error. By generating criteria for data correction, we developed a robust method to enhance data quality further. The application to a large set of experimental data validated our approach. The normalisation criterion not only validates normalisation procedures, but also assesses the quality of the immunoblotting experiment. Valid normalisation criteria and decreased standard deviation to the first estimate are indicators of adequate measurements. If our standard operating procedures for experimental design and data processing are employed, many aspects of quantitative immunoblotting can be automated.

Subtle parameter changes in biological systems can change the state of a cell and trigger the onset of diseases. Therefore quantitative measurements with the highest resolution possible are necessary so that we can understand, predict and interfere with these networks. The limitation of current systems biology is often the lack of data to test the quantitative accuracy of mathematical models, requiring new measurement techniques [4]. With the presented procedure, the established technique of quantitative immunoblotting is developed into a robust and reliable method to generate high-quality quantitative data for systems biology.

9 Acknowledgments

The authors would like to thank Nils Blüthgen for stimulating discussions and Peter J. Nickel and Verena Becker for reading the manuscript critically. We would also like to thank the nine undergraduate students participating in the practical course ‘Components and mechanisms of signal transduction’ for the data analysed in Section 7.

10 References

1 Swameye, I., Muller, T.G., Timmer, J., Sandra, O., and Klingmuller, U.: ‘Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by data-based modeling’, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, (3), pp. 1028–1033

2 Pristley, M.: ‘Spectral analysis and time series’ (Academic Press, London, 1989)

3 Wood, S.N.: ‘Thin plate regression splines’, *J. Royal Statistical Soc. B*, 2003, **65**, (1), pp. 95–114

4 Bhalla, U.S., Ram, P.T., and Iyengar, R.: ‘MAP kinase phosphatase as a locus of flexibility in a mitogen-activated protein kinase signaling network’, *Science*, 2002, **297**, (5583), pp. 1018–1023