BIOMARKER signal definition

Due to the numerous challenges in the industry, pharmaceutical companies are examining all aspects of the drug development process and rebuilding their associated business models as necessary. Companies need to improve success in late stage development, but what opportunities exist for more appropriately identifying the drug candidate to take to Phase III clinical trials?

By Dr Gordon F. Kapke and Dr Nigel Brown

Great emphasis is being placed on the development of new biomarkers through the FDA and the Critical Path Institute. The general consensus is that better biomarkers are required. Are there opportunities to improve the value of existing biomarkers through alternative data analysis processes? Should the same effort being placed on biomarker discovery be placed on biomarker data interpretation and presentation?

The identification of disease in the healthcare environment frequently depends on the reference interval to identify unusual results. The clinical chemist recognises the reference interval is of limited utility, and understands the Index of Individuality (IOI) varies across individual analytes. The IOI is the ratio of the intra-individual biological variation divided by the group variation. The IOI ranges from a low of 0.36 for serum creatinine to 1.03 for serum potassium. When the IOI is less than 0.6, the individual's results will span a limited portion of the reference interval. Subsequently, the reference interval can be of limited value in determining if a change has occurred within a subject. From a diagnostic perspective, the healthcare provider typically does not have a baseline value. Consequently, in the healthcare setting, the patient cannot be evaluated against himself or herself upon presentation to the physician, but must have some reference point, which is the reference interval.

Traditional data analysis looks at statistical significant variation by cohort, the relationship of the obtained values to the reference interval, or a change from baseline based on percent changes or absolute changes. In pre-clinical experiments with a very low number of observations, such as n = 5, differences between cohorts need to be large to detect a significant change for a biomarker. The statistical power of these pre-clinical experiments to identify small, but important biomarker changes is low. The low statistical power of small experiments contributes to the observation that pre-clinical experiments frequently do not translate into clinical observations. The data interpretation challenge is to find a process to define an important reliable signal when the observations are equal to 1 and the potential signal is small. Due to species differences, pre-clinical biomarker signals may be small. Phase I biomarker signals may also be weak due to the nature of the selected healthy population. Consequently, an improved process to identify weak signals in small clinical experiments may benefit drug development.

Serial results theory

Since the early 1990s, the clinical chemist has clearly understood how to define a significant change in a subject's serial results. The derivation of significant change in serial clinical chemistry measurements is clearly summarised in the book by Callum Fraser, *Biological Variation, from Principle to Practice* (AACC Press, 2001)[1]. In Chapter 3, Dr

Table 1: Intra-individual coefficient of variation (CV_i) by analyte. This table contains published CV_i for ALT, AST and Creatinine from two sources. A substantial difference in CV_i is found across analytes. The CV_is are in substantial agreement from the two sources

ANALYTE	% CV_i[1]	% CV_i[2]	IOI[2]
ALT	24.3	23.7	0.47
AST	11.9	15.1	0.52
CRT	5.3	6.8	0.36

IOI = CV_i / CV_g
CV_i = intra-individual CV%
CV_g = group CV%

1. www.westgard.com/biodatabase1.htm
2. Clin. Chem. 51. 450-452 (2005)

Fraser derives the theoretical approach to changes in serial results on a single patient.

The rationale for using the serial result theory for drug development data interpretation is that the goal in clinical experiments is to understand how a drug impacts a biological system. Therefore, a change from baseline in a pre-clinical study or in a human clinical trial is precisely the question that needs answering. Potentially, defining signal detection limits based on the theoretical limits generated from biological variation data may provide a means to identify important signals early in drug development and thus optimise candidate selection and minimise toxicity surprises in later stages of development. Biological variation defined Z values are significant before one reaches a limit like 2 x ULN and permits the identification of a significant reproducible change on every individual subject. The calculation of the Z value is the equivalent of personalised medicine on each subject. The definition of significant change can be applied and evaluated for each individual subject rather than the group level. Knowing that toxicities may be idiosyncratic, sensitive signal detection for each individual should permit identification of previously unrecognised biomarker signals even when the finding affects only a small part of the population.

The theoretical basis for identifying changes in serial patient results in the following equation:

$$Z = \frac{[(\text{visit} - \text{baseline})/\ \text{baseline}]\ 100}{\{2^{0.5}\ (\%CV_a^2 + \%CV_i^2)^{0.5}\}}$$

baseline = subject baseline biomarker value
visit = subject visit biomarker value
$\%CV_a$ = analytical imprecision
$\%CV_i$ = within subject biological variation for the biomarker

The details of the derivation of this equation are found in Chapter 3 of Dr Fraser's book.

Historically, the clinical chemist calculated Z to determine if Z exceeded a predetermined value associated with a confidence interval. In drug development, our proposal is to calculate Z from a data set and look to identify significant changes that may be unrecognised using traditional data analysis methods. The calculation of Z values appears to be more sensitive than historical data analysis methods to appropriately identify a significant reproducible change in an individual's clinical laboratory result.

Signal identification is critical in early development experiments and the Z value calculation may find information that was previously unrecognised. The observation of a signal may or may not be of clinical significance. Identifying a signal early, however, permits the opportunity to monitor and manage this potential effect before the commitment to a large expensive clinical trial.

Why hasn't the theory been used?

Since the theory for identifying a significant change in serial results has been known for 20 years, why hasn't this methodology been previously applied in drug development? There are assumptions that apply to this serial data analysis approach and there may be concern of the validity of the assumptions required for this calculation.

First, is the pre-analytical variation (CV_p%) truly small? Careless practices by a site can invalidate this assumption quickly. Our experience in plotting data obtained in the clinical development setting indicates that this is a reasonable assumption. Data analysis across random data sets has not generated inappropriate findings concerning spuriously low or high Z values, indicating that the pre-analytical phase of sample collection and handling is adequately controlled.

Second, CV_i data has been consolidated and published only relatively recently. CV_i data was first consolidated and published in *The Scandinavian Journal of Clinical Laboratory Investigation* in 1999[2] and additional CV_i data was published in *Clinical Chemistry* in 2005[3]. The published CV_i data agree appropriately (**Table 1**). The most complete, easily accessible database for CV_i values can be found at www.westgard.com/biodatabase1.htm which is updated annually. Although the CV_is are not identical in the two databases, the material difference between the two databases is viewed to be insignificant in the calculation of Z values compared to the differences in

Z values signal definition relative to tradition alert limits.

Third, is the variation of CV_i appropriately small such that when analysing data one observes appropriate Z value distributions? There is no direct manner to determine the consistency of CV_i across populations, over time, geography and clinical condition, however, using published CV_is, calculation of Z values on global data suggests that CV_i is appropriate across these four areas.

The variability in CV_i by analyte is not widely recognised. Intuitively, one would expect CV_i to be different for different analytes. But, what is the range of CV_i across analytes? Which analytes have high CV_is and which analytes have low CV_is? **Table 1** has a list of three analytes and published $\%CV_i$s. Note that the serum creatinine $\%CV_i$ is 5.3% and the ALT $\%CV_i$ is 24.3%. CV_i has a considerable range and thus, changes beyond biological variability will be significantly different across analytes.

The serial results model shows that the CV_i does directly affect the level of a significant change in analyte concentration. **Figure 1** is a graph of a family of lines that were calculated for multiple CV_is with a constant CV_a of 5%. The CV_i values were chosen to cover the range found across many clinical analytes. CV_i values were chosen at 5, 10, 20, and 30%. Given the CV_i and CV_a, the Z values were calculated for increases in analyte concentration. Note that with a CV_i of 5%, which is very similar to the creatinine CV_i, a 30% increase in serum creatinine concentration is an important change. For a CV_i of 10%, which is similar to the AST CV_i, an increase of 45% is an important change. For a CV_i of 20%, which is similar to the ALT CV_i %, a change of 90% from baseline is an important change. An important change may or may not be clinically significant. Given the small sample size in preclinical and early clinical development, the calculation of the Z value provides a sensitive signal detection methodology for each discrete observation and will identify signals that would not be recognised using group statistics.

Finally, what is the appropriate change in Z to determine a change that defines an important signal? From a confidence interval perspective, a Z change of 1.96 represents a 95% confidence interval and a Z change of 2.58 indicates a 99% confidence interval. Due to questions about the variability in the assumptions, one may question the use of the theoretical 99% confidence interval. Our experience has been that a Z value of 3 should be considered an important signal.

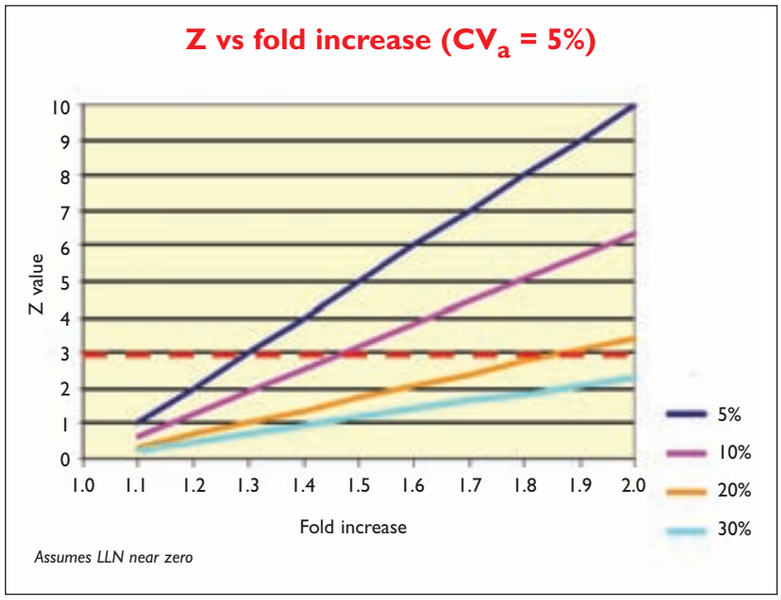

Figure 1: A family of lines has been calculated and graphed for multiple CV_i values using a constant CV_a of 5%. Assuming a Z value of 3 defines the significant signal, the important fold increase varies from 30% to 90% as CV_i moves from 5% to 20%

Data analysis examples

In reviewing a creatinine data set, the data has interesting observations. In **Table 2**, 5% was used for the analytical imprecision (CV_a) and the CV_a was taken from the laboratory QC documents for the lowest control, and therefore

Sample creatinine Z values

BASELINE CREATININE µMOL/L	VISIT CREATININE µMOL/L	DELTA CREATININE µMOL/L	Z
56.6	91.9	35.3	6.6
92.8	133.5	40.7	4.7
59.2	82.2	23.0	4.1
71.6	91.0	19.4	2.9

$$Z = \frac{\left\{\left(\frac{91.9 - 56.6}{56.6}\right) \times 100\right\}}{2^{0.5} \times (4.0^2 + 5.3^2)^{0.5}}$$

CV_a = 4.0% (from lab QC lowest control)
CV_i = 5.3% (from www.westgard.com/biodatabase1.htm)

ULN = 110 µmol/L

Table 2 shows baseline and visit results for selected patients. All baseline values are within the reference interval. There is a poor relationship among visit value, delta value and Z value. Calculating the change from baseline using Z value may improve biomarker signal detection

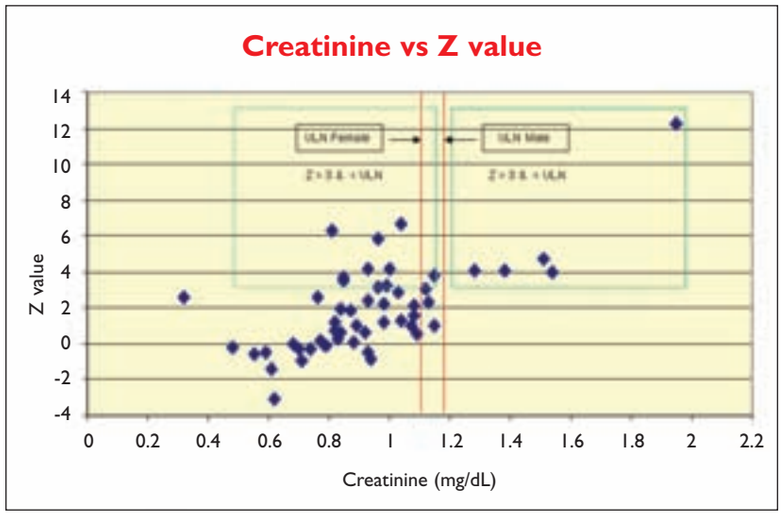

Figure 2: Visit value is graphed against the calculated Z value for the study population. The Z value can be significant in the reference interval and may be insignificant above the reference interval

represents a relatively high CV_a. The CV_a will vary with analyte concentration so it is important to take this value from the control data that is in the range of the majority of the clinical data.

The first subject started with a baseline value of 56.6 µmol/L and at the defined visit has increased to 91.9 µmol/L. The change was 35.3µmol/L and the calculated Z value was 6.6 and was very significant. The second subject had a baseline value of 92.8 µmol/L and a visit value of 133.5 µmol/L (reference interval ULN being 110 µmol/L). The change was 40.7 µmol/L and the calculated Z value was 4.7 and significant. The third subject had a baseline value of 59.2 µmol/L and a visit value of 82.2 µmol/L for a change from baseline of 23.0 µmol/L and a calculated Z of 4.1, which is again significant. The last subject had a baseline value of 71.6 µmol/L and a visit value of 91.0 µmol/L. The calculated Z value was 2.9, which is marginally significant. Note that the magnitude of the delta in the visit value versus the baseline value does not correlate well with the calculated Z value. Consequently, in traditional analysis of data, one may have thought they had an important signal when the change was within expected biological variation and one may have ignored changes within and around the reference interval as not being important even though the Z values were significant.

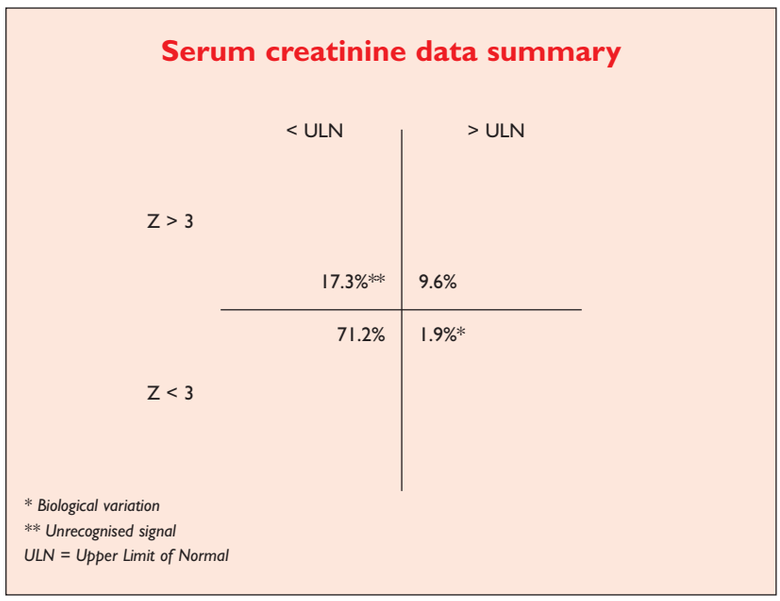

Figure 3: An alternate presentation of the data is to define one axis as a traditional limit (> or < ULN) and the second axis as the Z value signal limit. The traditional limit may miss Z value defined signal. The low IOI for creatinine predicts that signal will be missed within the reference interval and the data confirms this phenomenon

The second aspect of looking at data from a Z value perspective that is unique is that the Z value equation is method and reference interval independent. Therefore, across a clinical trial population that has been tested in multiple laboratories with multiple methods, a signal detection process can be defined to eliminate the differences in methods and the differences in local reference ranges. Due to the method and reference interval independence of Z value calculations, the use of Z values in analysing local laboratory data may provide more effective detection of a safety signal in local laboratory-generated data.

Plotting the data of visit value versus the Z value provides a perspective where important changes below traditional alert limits (Z > 3 and result < ULN or < 2xULN) are easily recognised, and increased values outside traditional limits can be recognised as being associated with normal biological variability (Z < 3 and result > ULN).

Figure 2 is the graph of serum creatinine visit value versus Z values for a small data set. All values above the reference interval have Z values greater than 3. A greater number of results within the reference interval also had Z values greater than 3. If the results outside the reference intervals are significant and have Z values greater than 3, the values within the reference interval with Z greater than 3 should be evaluated with equal significance.

The data in **Figure 2** can be visualised as a four-quadrant graph (**Figure 3**). Counting values with Z > 3 and partitioning the values above and below the reference interval and the values with Z < 3 and above the reference interval gives a table as shown in **Figure 3**. Note that more values with a Z > 3 are

below the reference interval than above. The values with Z < 3 and greater than the reference interval are referred to as biological noise. Values with Z > 3 and within the reference interval are referred to as unrecognised signals. Perhaps viewing signals as all values with Z > 3 would result in different data interpretation than viewing the values greater than the upper limit of normal. The observed signal is also much stronger based on the Z values than based on the upper limit of normal. The increased signal intensity generated by including all Z values greater than 3 is directly related to the small IOI for serum creatinine.

Alternatively, converting the data in **Figure 2** to a histogram generates the graph in **Figure 4**. The histogram is not symmetrical and the data is weighted to increased Z values. This is an unusual distribution for serum creatinine.

Figure 5 is the histogram of serum creatinine data from a Phase I clinical trial. The data is evenly distributed around a Z of zero and decreases as expected as the absolute value of Z increases. This symmetrical Gaussian-like distribution is typical of the absence of a biomarker signal.

Moving into the Phase II population, does this nice symmetrical distribution of Z values observable across visits and over time persist for a patient population? **Figure 6** is an example of the Z value histograms observed over multiple visits on a patient population. Note that the Z value distribution of serum creatinine remains quite symmetrical and evenly distributed over expected Z values for the multiple visits. In the absence of a creatinine signal, the data distributions remain very similar over time.

The symmetrical Z value histogram on the Phase I population and multiple symmetrical histograms on the Phase II population indicate that the non-symmetrical histogram distribution in **Figure 4** is an unusual observation. The non-symmetrical shape and right shift of the Z values distribution in **Figure 4** are consistent with a weak biomarker signal.

The display of the data as a plot or a histogram has the added benefit of providing all reviewers a clear consistent picture of how the data is distributed. In statistical summaries, important changes may be unrecognised. By reviewing a Z value graph or histogram, a clear picture of how data is changing over time for a specific biomarker is available to all data reviewers. A clear and consistent picture of all the data for all reviewing parties may permit debate that generates optimal development decisions.

Values for the definition of important biomarker changes have a theoretical basis using serial result analysis theory. The significance of the biomarker value changes relatively to multiples of the upper reference limit (URL) or fold increase in analyte concentration are dependent on the analyte's CV_i. Use of serial result theory provides a strong theoretical framework for identifying important changes and a bases to set limits from measured parameters (CV_i and CV_a).

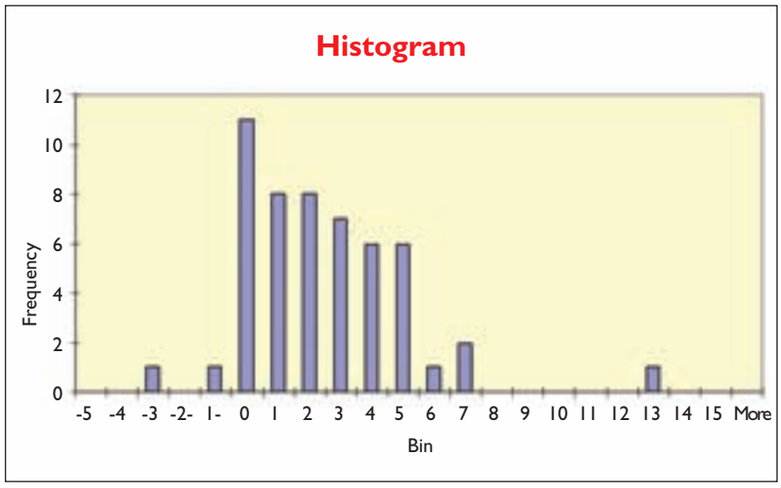

Figure 4: A third presentation of data is a histogram. The histogram has lost symmetry and the median of the histogram shifted significantly to the right. The appearance of this histogram is consistent with a weak biomarker signal

Can archived data sets be mined using this technique?

Sponsors have development data on many compounds. Serial results theory provides a tool to mine existing data. The Z value analysis is independent of the reference interval, and also independent of the method. The Z value calculation depends only on the percentage change from baseline and the CV_i and CV_a. CV_a can be estimated from published data and the dominant determining factor is the CV_i. The CV_a is typically smaller or

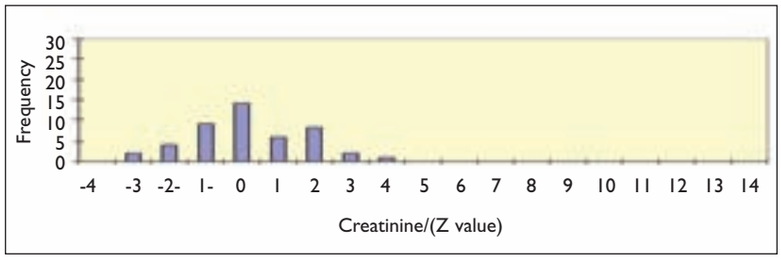

Figure 5: The histogram of serum creatinine Z values is from short-term experiment on a small population. The histogram is centred on a Z value of zero and is symmetrically distributed. These are the observations for a typical data set with no signal present

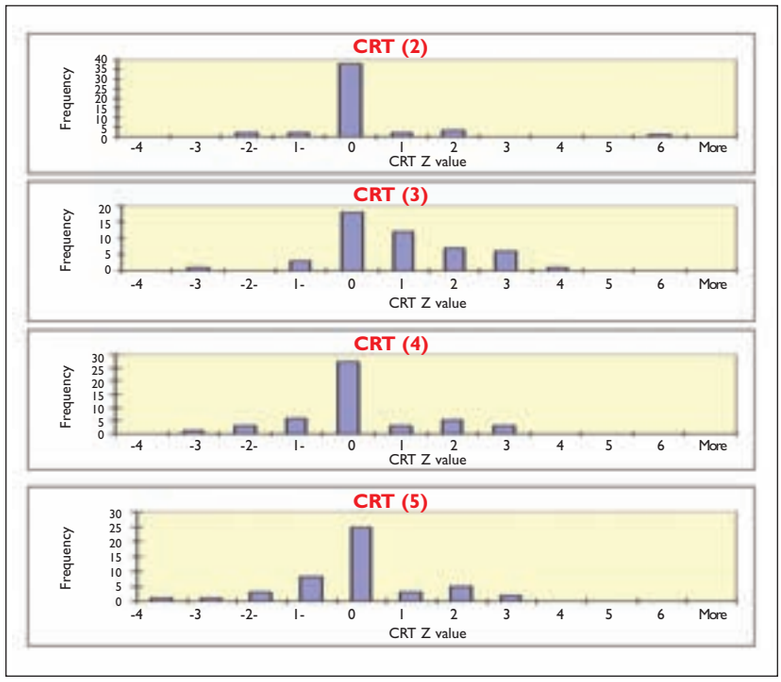

Figure 6: The histograms of serum creatinine Z values is from multiple visits on a small population. Over multiple visits there are some minor changes in Z value distributions. However, the Z values are nicely centred around Z = 0 and the distribution remains symmetrical over the five visits

References

1 Fraser, Callum G. Biological Variation, from Principle to Practice, AACC Press (2001), 2101 L Street NW, Suite 202, Washington, DC 20037-1558.
2 Ricos, C, Alvarez, V, Cava, F, Garcia-Lario, JV, Hernandez, A, Jimenez, CV, Minchinela, J, Perich, C and Simon, M. Scan J of Clin Lab Invest 59: 491-500, 1999.
3 Lacher, DA, Hughes, JP and Carroll, MD. Clin Chem 51: 450-452, 2005.

equal to the CV_i so the CV_i is the major determinant in defining a significant change. Important changes happen before one crosses traditional fixed limits like ULN or 2 times ULN.

In order to have confidence that the transformation of data into Z values is beneficial, examples of how development decisions might have been affected retrospectively needs to be published. Drugs that have failed development in late stages should be retrospectively reviewed by Z value analysis for selected biomarkers so the industry can understand how an unsafe drug's Z values histograms evolved along the development pathway. Only with concrete examples in the literature will this approach to data analysis be embraced.

Summary

In summary, Z value analysis is based on measured parameters to include the change from baseline, the CV_a, and CV_i. The CV_a and CV_i are measured parameters that can be continually challenged and confirmed. Significant Z values will recognise smaller biomarker result changes than traditional limits like 2 x ULN. The Z value methodology is independent of reference interval and analytical method, therefore historical data can be analysed to understand how important signals have developed during discontinued projects. DDW

Dr Gordon Kapke *is Senior Director of Biomarker Services for Covance Central Laboratory Services (CCLS). Dr Kapke joined CCLS in 1991 as a Laboratory Technical Director and has made key contributions including design and implementation of CCLS laboratories in Geneva, Sydney, Singapore and China. In China, Dr Kapke also coached Huashon Hospital Clinical Laboratory in Shanghai to CAP accreditation. His scientific and laboratory operational experience has been instrumental in making Covance one of the world's leading central laboratories. Dr Kapke holds a BS degree in Chemistry from the University of Wisconsin and a PhD in Chemistry from the University of Iowa. After graduate school, he accepted a position as a research scientist in the Department of Pharmacology, University of Iowa, followed by a clinical chemistry post-doctoral fellowship in the Department of Pathology, University of Iowa. Dr Kapke is a Diplomat of the American Board of Clinical Chemistry and a Fellow, National Academy of Clinical Biochemistry.*

Dr Nigel Brown *received his bachelor's degree (chemistry) and his doctorate (natural products metabolism) from the University of Oxford, UK, funded by the US National Institutes of Health. He was subsequently awarded a Nestlé postdoctoral fellowship at Massachusetts Institute of Technology, where his research focused on the cancer-forming potential of several common food products. He joined the Department of Drug Metabolism and Pharmacokinetics at The Upjohn Company in 1991, where he supported drug metabolism research and first-in-human clinical trials. In 1993 Dr Brown moved to SmithKline Beecham R&D where he played a key role in introducing LC/MS/MS as the principal analytical tool for metabolic and pharmacokinetic assessment. In 1997 he joined Phoenix International Life Sciences, a high-growth CRO, in a general management capacity overseeing preclinical operations. Following the acquisition of Phoenix International by MDS, Dr Brown was General Manager of MDS Pharma Service's global preclinical enterprise. After receiving his MBA, Dr Brown transitioned his focus to Corporate Strategy and M&A within the CRO sector and has subsequently built a solid deal list. He has worked for Covance since June 2006 as head of corporate development and strategy. Dr Brown sits on the management team of Covance Inc, serves on the boards of several healthcare companies, and has published broadly in scientific and business journals.*