Examination of Some DNA Guidelines Using Simulation Data

Donald J. Herbert, C. Bruce Bagwell
Mark E. Munson and Benjamin C. Hunsberger
Verity Software House, Inc., PO Box 247, Topsham, ME 04086 USA

The "Guidelines for Implementation of Clinical DNA Cytometry" published in 1993 established a number of criteria for DNA ploidy and S-phase analysis. This study examines several of those criteria to determine their relative importance in accurately measuring S-phase percentage.

We examine the following criteria and their effect on errors made in determining S-phase percentage:

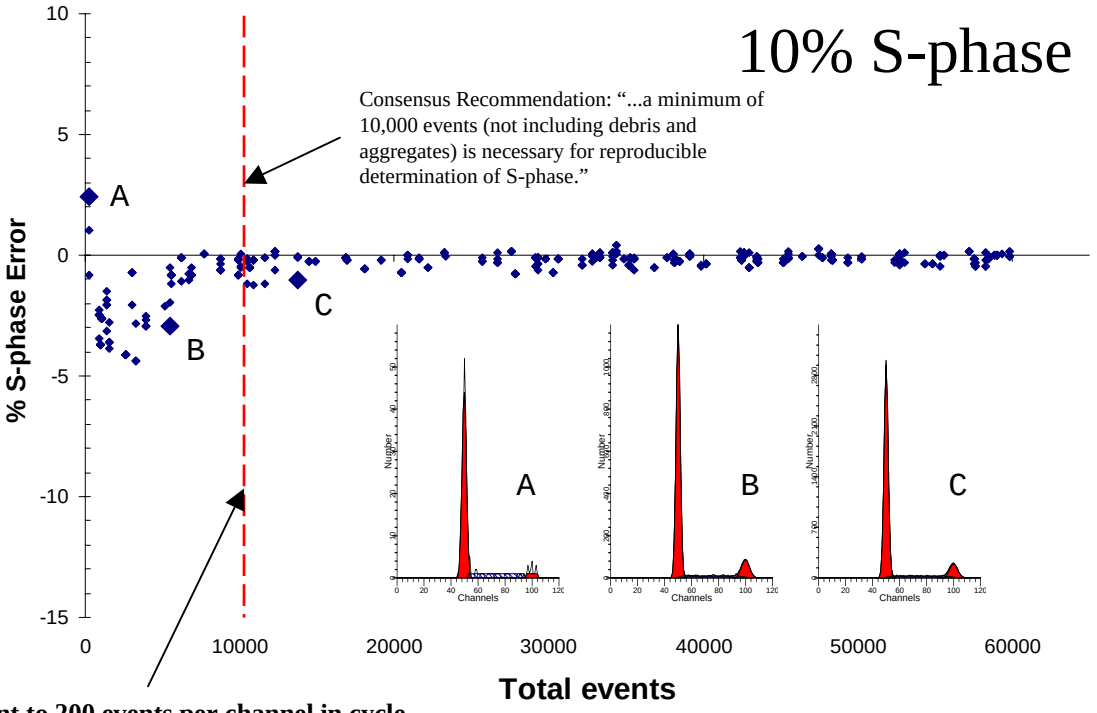
1. Number of events
2. Number of channels
3. %CV of populations
4. Aneuploid fraction
5. Debris and aggregation (% BAD)
6. DNA Index (DI)
7. Reduced Chi-Square (RCS)

For each criterion, we use computer software to synthesize a minimum of 500 histograms of defined composition, varying only the selected criterion. Each histogram is analyzed using ModFit LT 2.0 DNA analysis software. The analysis is databased for error analysis using Microsoft Excel.

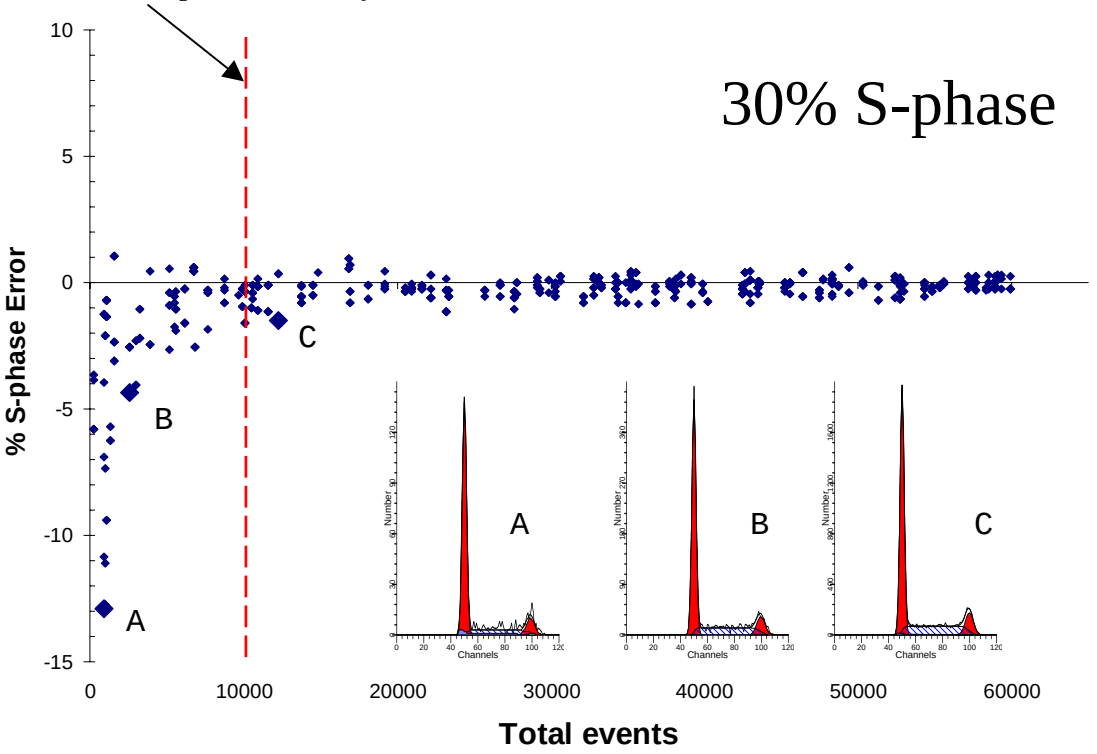
Results of the analysis are presented, some of which are quite surprising. We find that S-phase can be estimated with relatively few channels in the histogram, as long as the average number of events per channel is sufficient. Reduced Chi-Square is not a good measurement of quality for histograms with skewed peaks or shoulders. This study supports the DNA Consensus proposed value and use of %BAD as an exclusion criterion.

The authors gratefully acknowledge the participants of the 1997 Annual Research and Clinical Flow Cytometry Courses held at Los Alamos National Laboratory, NM and Dartmouth-Hitchcock Medical Center, NH for their contributions to this study (see participant list below).

1. Number of Events

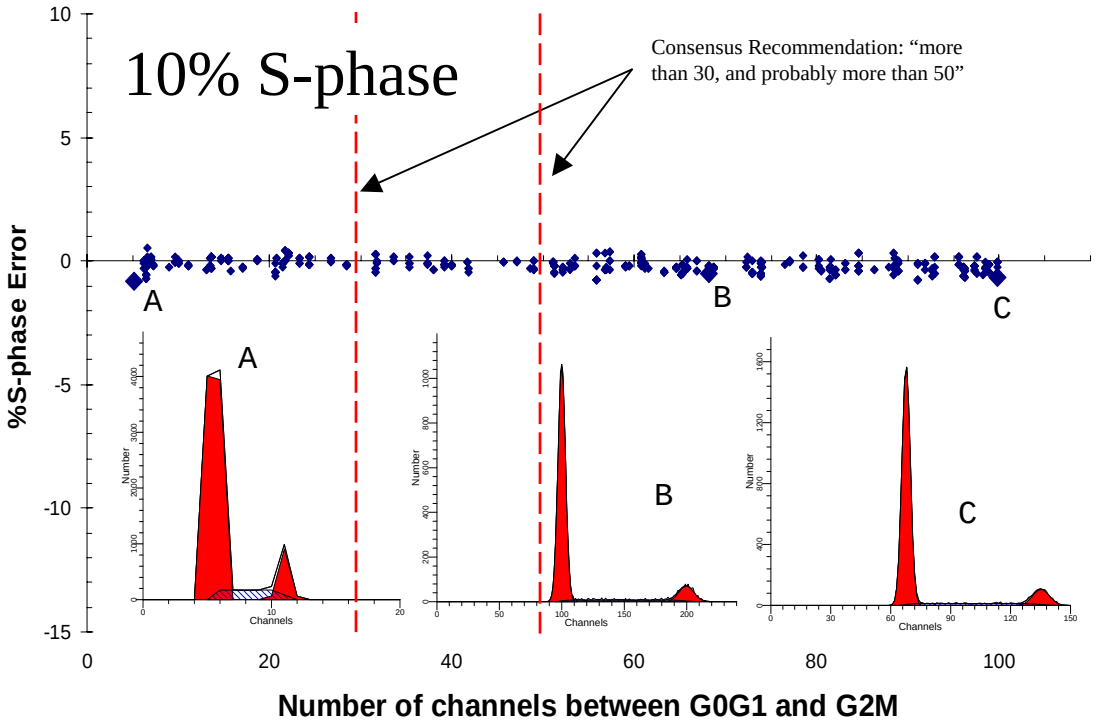


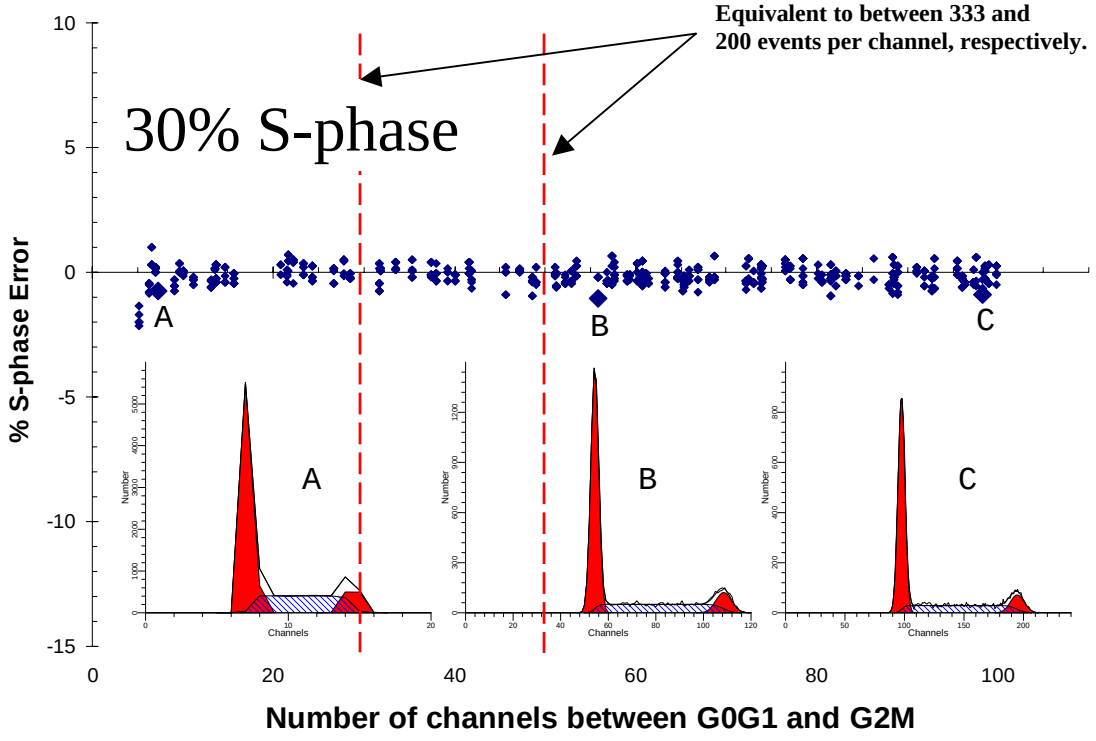
Equivalent to 200 events per channel in cycle.



These data support the DNA Consensus recommendations for collecting 10,000 events (Equivalent to 200 events per channel in cycle) for analysis of S-phase.

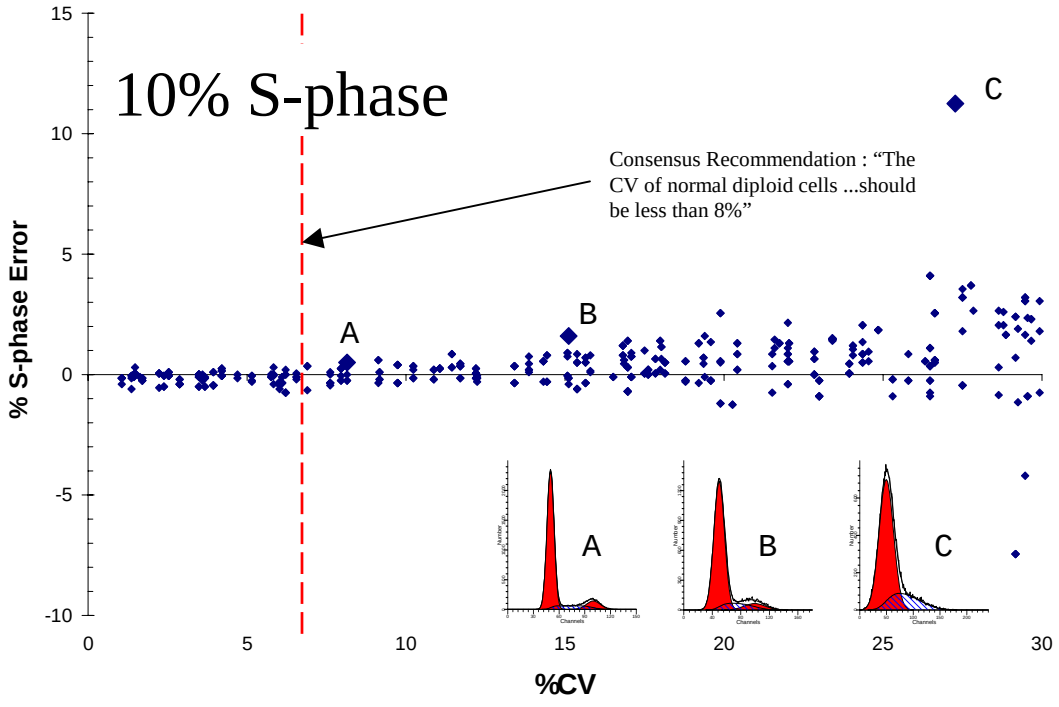
2. Number of Channels

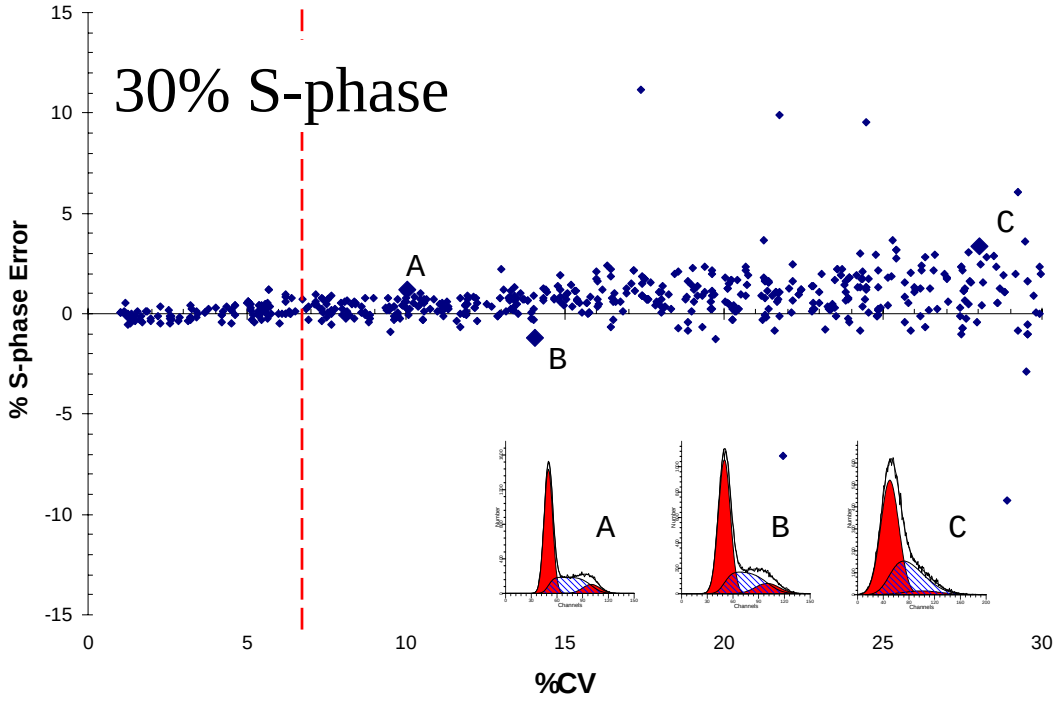




The Consensus Conference suggested that analysis should be done with 30 or more channels. These data show that good S-phase results can be obtained with a small number of channels.

3A. %CV
DNA Diploid

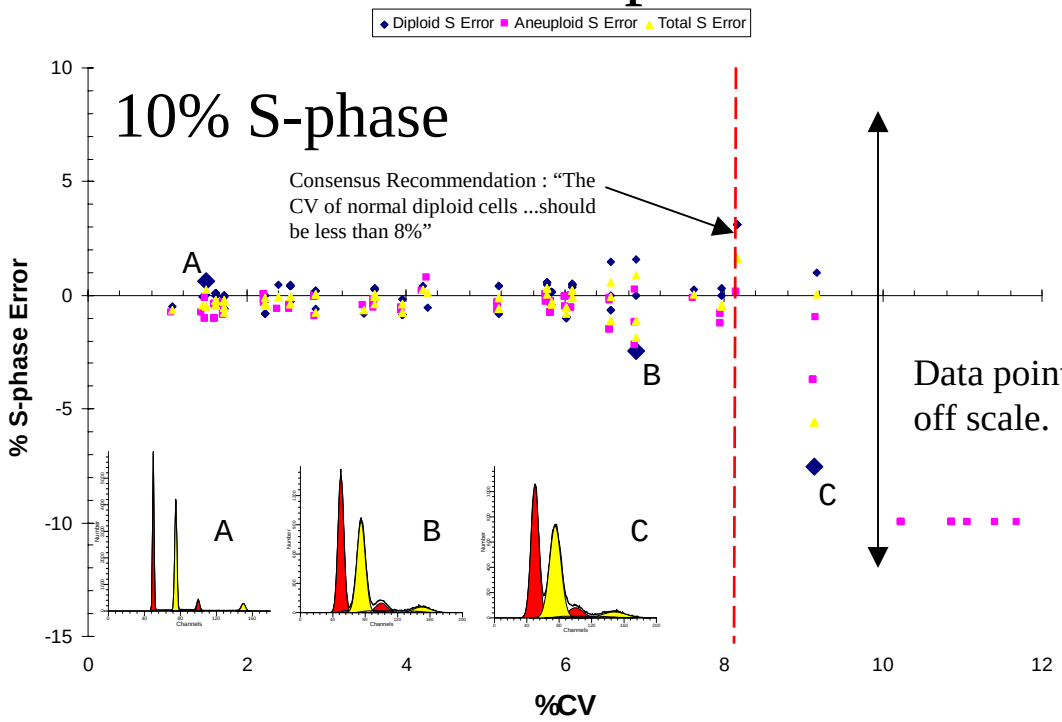


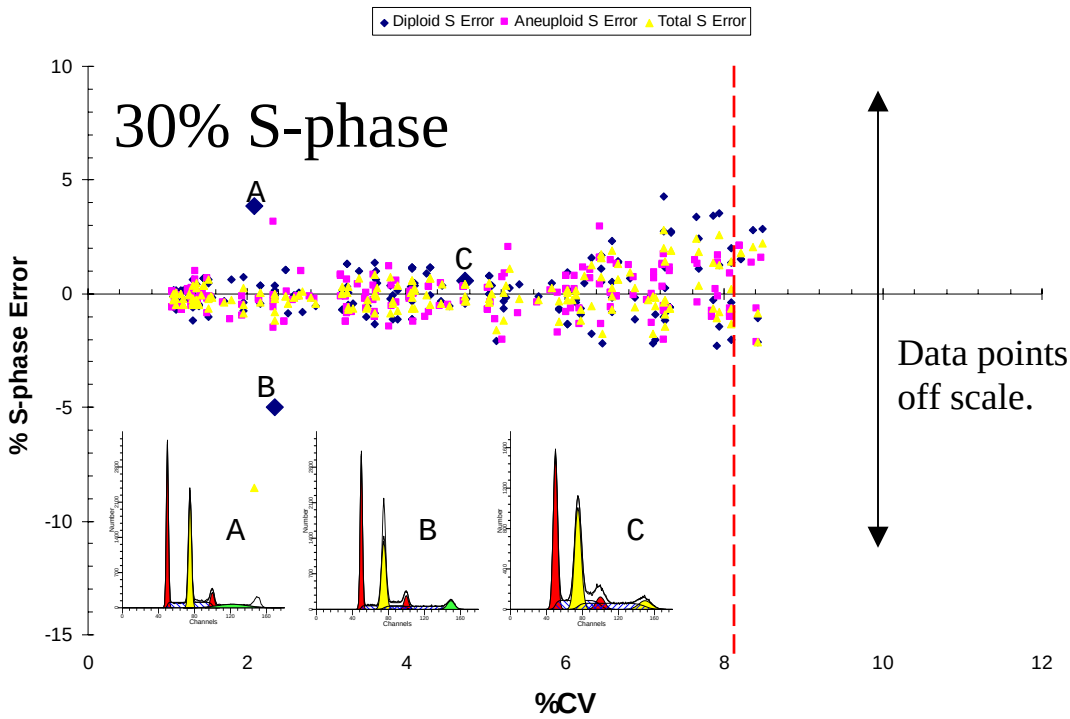


These data show that DNA diploid samples can be accurately analyzed with %CV's above the DNA Consensus recommendations.

3B. %CV

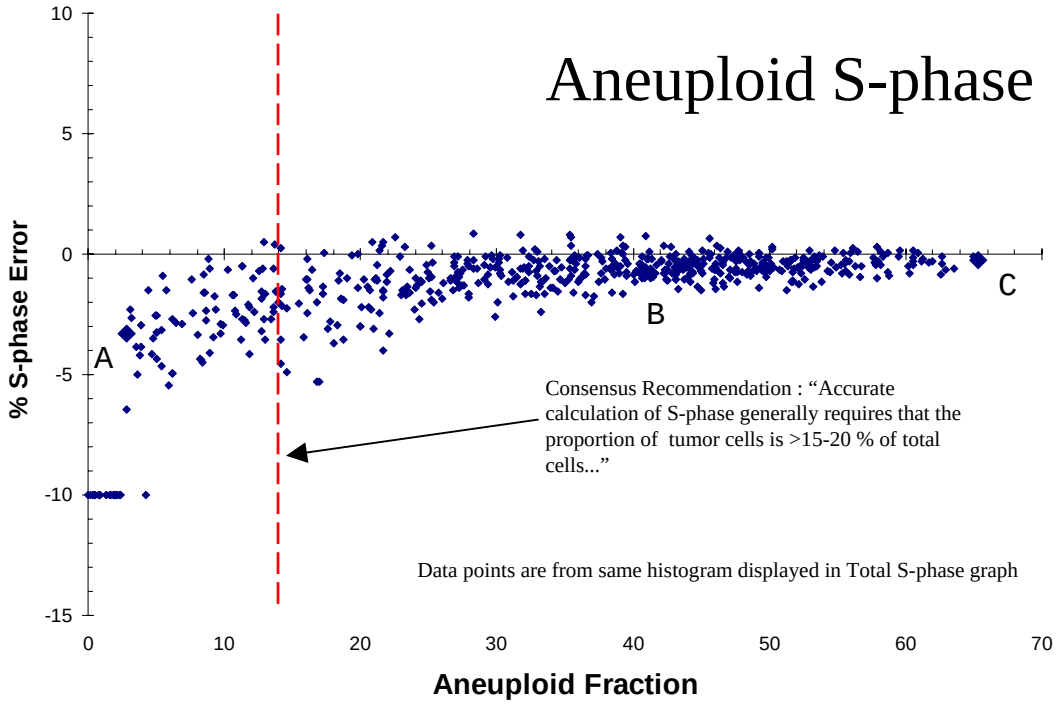
DNA Aneuploid

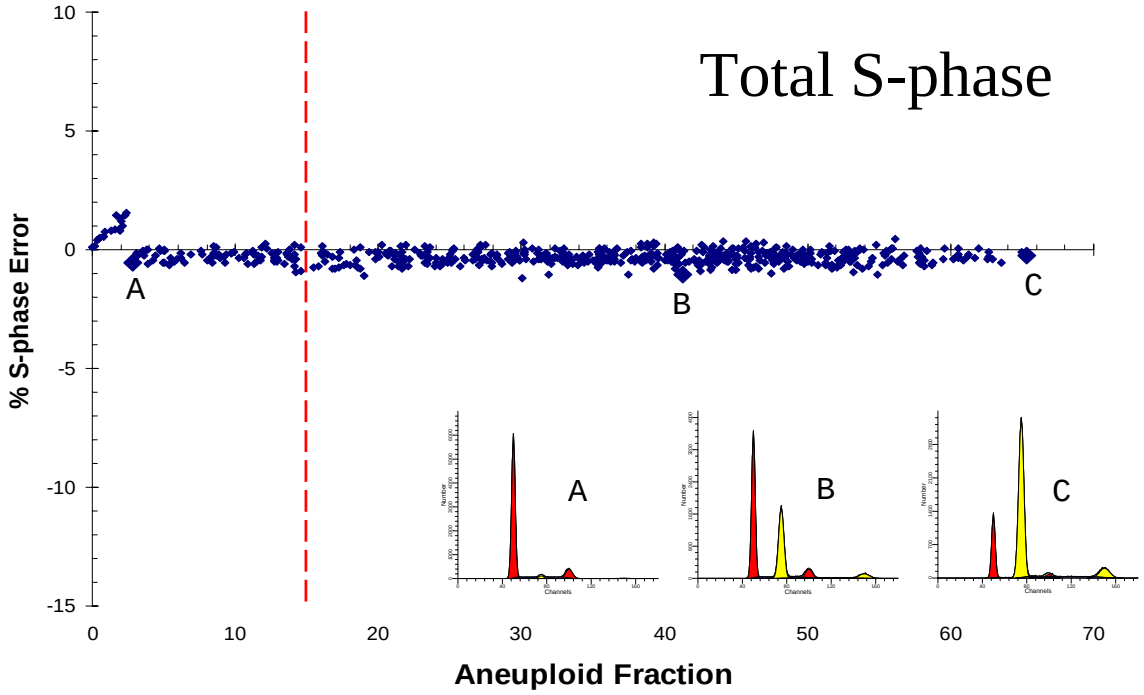




These data show S-phase error for Diploid, Aneuploid, and Total populations generally increase with %CV. Above 8% CV, S-phase values were artifactually off scale due to the software's inability to discriminate peaks for automatic analysis.

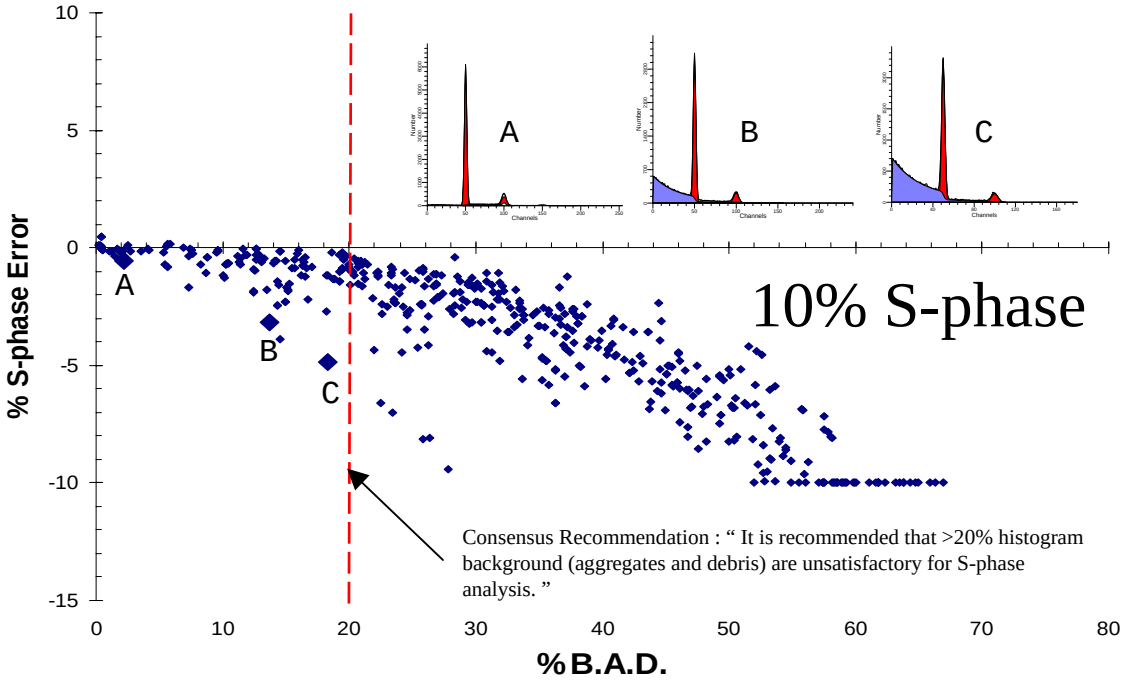
4. Aneuploid Fraction

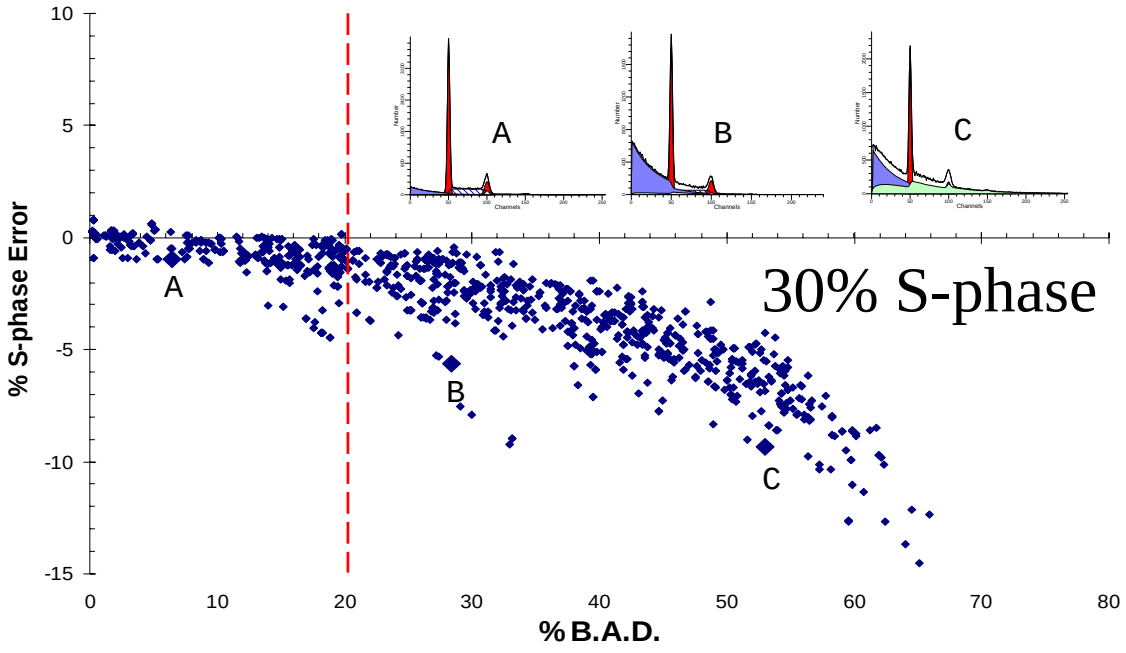




These data show that Total S-phase values are generally more accurate than Aneuploid S-phase especially for low aneuploid fractions.

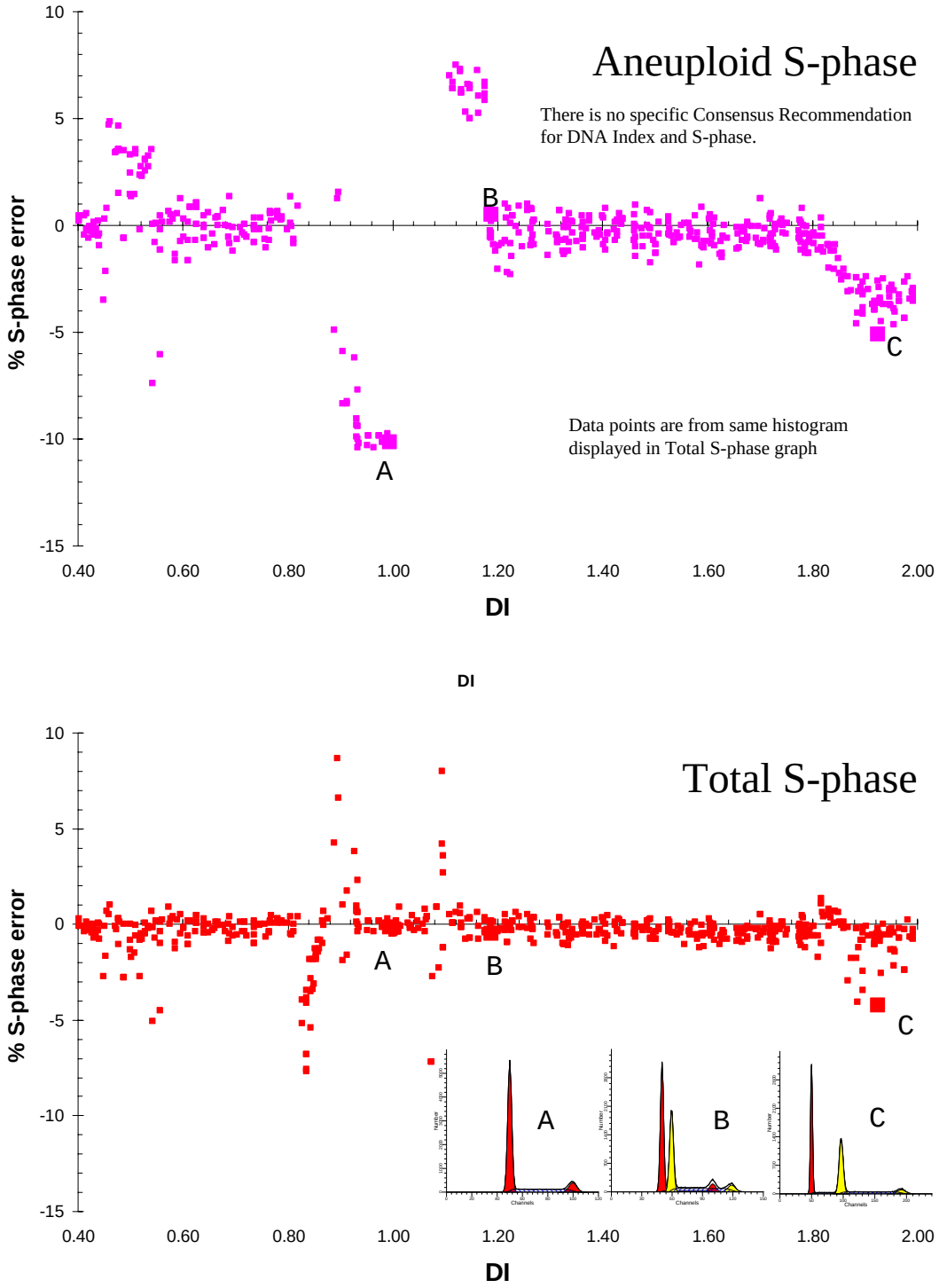
5. % B.A.D.





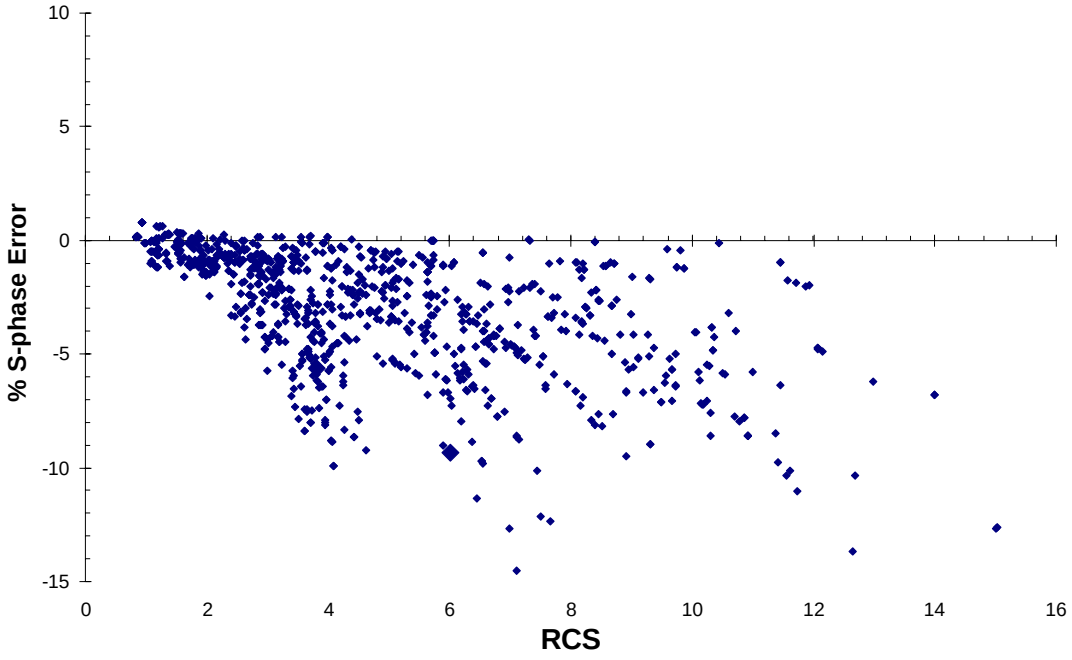
These data show that S-phase error increases as %B.A.D increases.

6. DNA Index (DI)



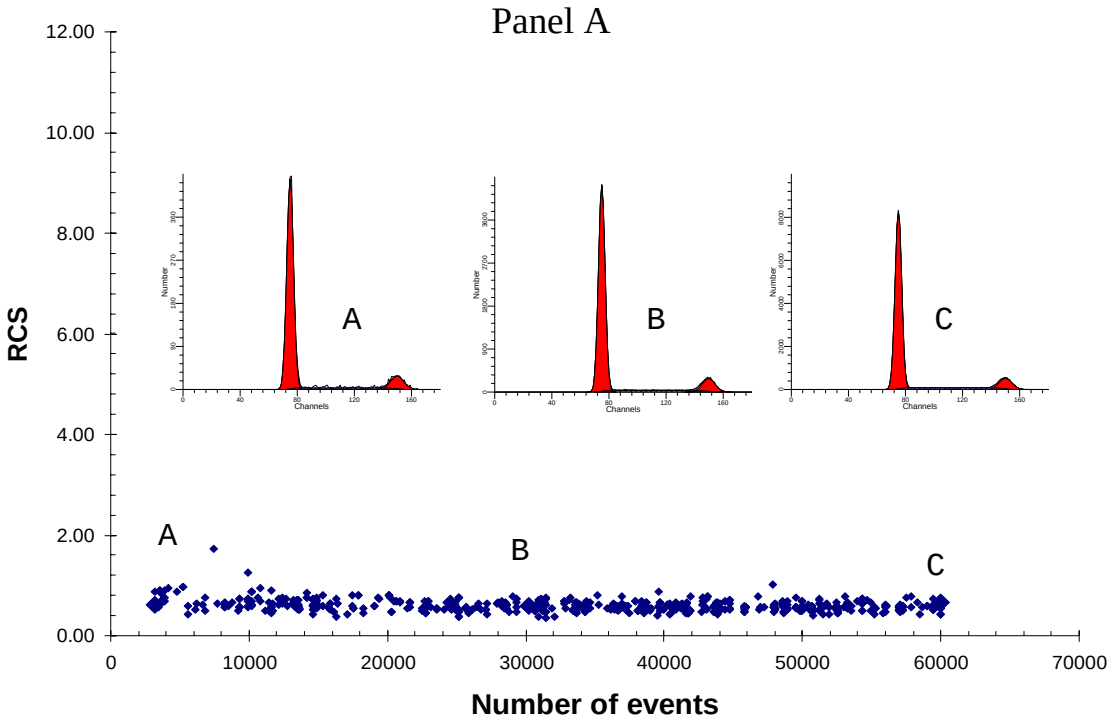
These data show that Aneuploid S-phase results can not be accurately obtained between DNA indices of 0.8 and 1.09 or between DNA indices 1.81 and 2.0. As is also shown in Figure 4, Total S-phase results are generally more accurate than Aneuploid S-phase.

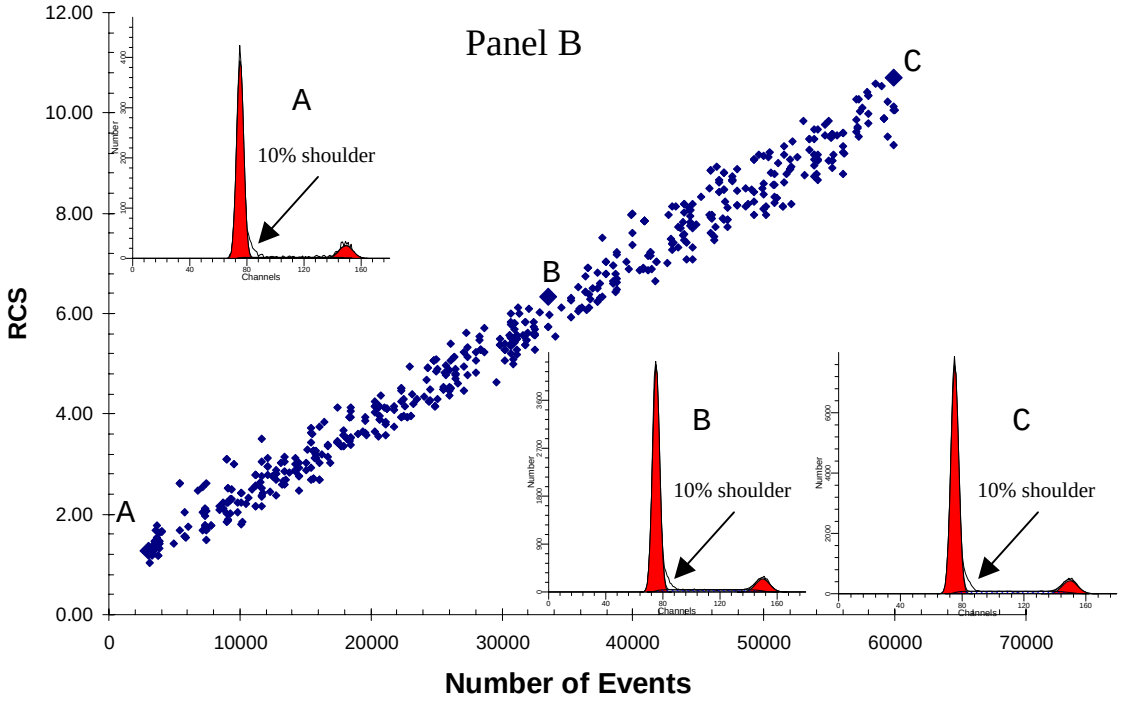
7A. Reduced Chi-Square (RCS)



This panel shows RCS and S-phase results from the data set used in the %B.A.D. Study. These data suggest a relationship between RCS and %S-phase error. As shown in the next panels, however, RCS may not directly relate to % S-phase error.

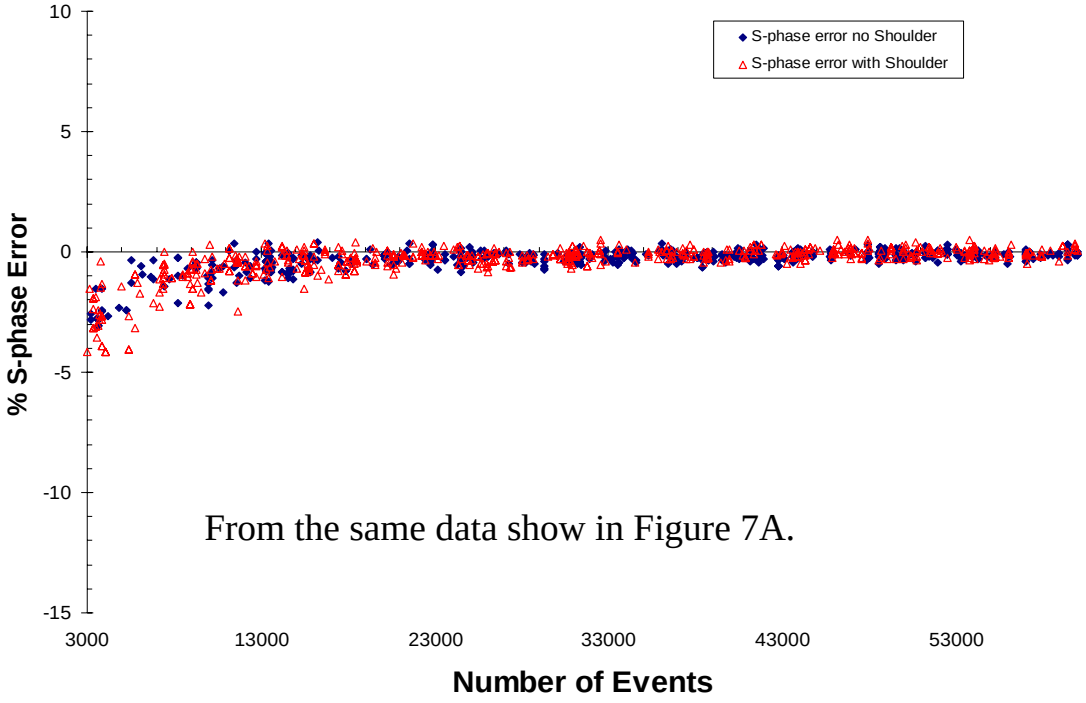
7B. Reduced Chi-Square (RCS)





Panel A shows that when the model is appropriate for the data, RCS values are close to 1.0. Panel B shows that when model and data do not match well, the RCS values linearly increase with number of events.

7C. Reduced Chi-Square (RCS)



From the same data show in Figure 7A.

This figure shows the S-phase error for histograms from both panels of Figure 7B. Both groups report accurate S-phase even though their RCS values differ widely.

Summary and Discussion

This study began as a way to show students at the annual clinical and research flow courses what factors might affect S-phase analysis. It quickly became evident that we should expand the work done by the students using larger data sets and present the work in a formal setting. The histogram generator used, TestDNA, is available to others via our Web page (www.vsh.com) to allow further simulation studies.

Our goal was to look at some DNA Consensus Conference guidelines (1) and two other parameters, DNA index and reduced chi-square (RCS), not addressed by the Consensus Conference document. The study design enabled us to examine individual factors associated with DNA histograms. To minimize operator subjectivity, all analyses were performed with ModFit LT 2.0, using automatic analysis with standard peak discriminator settings.

Number of events:
According to the data presented here, it is better to describe this criteria as average number of events per channel. If 50 channels are used to distribute the cell cycle events, then the guideline transforms to 200 events per channel. This is a better formulation of the guideline since it is independent of number of channels. According to this reformulation , data that is distributed over more that 50 channels requires a higher total number of events to maintain the 200 event per channel criteria. Although difficult to see in Figure 2 because of scaling, as the number of events per channel decrease to below 200, the S-phase error increases.

Number of Channels:
The Consensus Recommendation states that “more than 30, and probably more than 50” channels are required, this study revealed, however, that as few as 10 channels gave good S-phase estimates, even when only 10% S-phase was present in the histogram

% CV of DNA Diploid populations
In contrast to the Consensus Recommendation that “the CV of normal diploid cells… should be less than 8%,” this study found that diploid G0G1 cells with a %CV of up to 20% gave reliable S-phase values.

Summary and Discussion

% CV of DNA Aneuploid populations:
In a mixed diploid/aneuploid sample, however, the %CV has a profound effect on the ability to accurately model S-phase. Above the recommended 8% CV cutoff, diploid S-phase, aneuploid S-phase and total S-phase all became unreliable, supporting the original Consensus Recommendation guidelines. It should be noted that errors in S-phase for histograms showing a %CV of greater than 8% were artifactually off scale due to the software's inability to discriminate peaks using automatic analysis. An experienced operator using manual analysis may fare better than these results would indicate.

Aneuploid fraction:
A series of histograms in which the aneuploid fraction population was varied between 0% and 80% was analyzed to examine its effect on aneuploid S-phase and total S-phase. Our analysis revealed that for histograms whose aneuploid fraction fell below roughly 30%, the aneuploid S-phase estimate was unreliable. In this subset diploid S-phase estimates were generally acceptable. Perhaps not surprisingly, in histograms with an aneuploid fraction of roughly 80% or greater aneuploid S-phase estimates were found to be more accurate than the corresponding diploid S-phase estimates. These data suggest that the original Consensus Recommendation requiring that the "proportion of tumor cells (be) >15-20% of total cells" may be underestimating the cutoff. In contrast to diploid S-phase or aneuploid S-phase estimates, however, and most importantly, determination of total S-phase remained generally accurate over the entire range.

Debris and aggregation (% B.A.D.)
Aggregates and debris were concerns at the DNA Consensus Conference. For determining the effect of % B.A.D. (background aggregates and debris), diploid histograms were generated with both aggregates and debris allowed to co-vary over a range of 0-90%. In histograms containing generated 10% S-phase, S-phase error increased rapidly to unacceptable levels with %B.A.D. above 20%. This supports the original Consensus Recommendation "that >20% histogram background… (is) unsatisfactory for S-phase analysis." As the proportion of S-phase in the histogram increased, this 20% cutoff may be extended somewhat, but

(continued)

above roughly 20% B.A.D. error increases significantly. No attempt was made here to determine whether debris or aggregation had a more prominent role in this regard than the other.

DNA Index (DI)

Our data indicate that DI values between the ranges of 0.8 to 1.1 and 1.81 to 2.0 produce unreliable S-phase estimates. Studies done by course students that included varying both %CV, and DI suggested that with increasing %CV these ranges might be even wider.

Reduced Chi-Square (RCS)

Reduced Chi-Square measures the adequacy of a DNA model's fit to data and therefore has potential as an additional acceptance criteria not covered by the DNA Consensus Conference. The data presented in Figure 7A seems to support this supposition. Higher RCS values are generally associated with poorer S-phase estimates.

However, the data presented in Figure 7B -7C show an important exception to using RCS as an acceptance guideline. If the model does not completely describe the data, the RCS will increase linearly with number of events (see Figure 7B, Panel B). This linear increase is relatively independent of S-phase error (see Figure 7C). We simulated this problem with a peak shoulder that is not modeled, but the same problem can be observed with skewed peaks, or incomplete modeling of debris, aggregates, or S-phase. Because of RCS's sensitivity to these relatively commonplace problems it is better to use RCS as a measurement of model performance than as an acceptance criterion.

1. T. Vincent Shankey, Peter S. Rabinovitch, Bruce Bagwell, Kenneth D. Bauer, Ricardo E. Duque, David Hedley, Brian H. Mayall, and Leon Wheeless: Guidelines for Implementation of Clinical DNA Cytometry. Cytometry 14:442-447, 1993

Materials and Methods

Histograms were generated using TestDNA, a freeware program available from Verity Software House, Inc. (www.vsh.com). This program can synthesize multiple DNA histograms with known cell cycle compositions. Important parameters such as %CV and number of events can be randomly selected over a given range. The generated histograms were saved as Flow Cytometry Standard files.

All S-phase components are broadened rectangles. Once the criteria were selected, a minimum of 500 histograms were created. With the exception of %B.A.D., only one criterion was varied in each experiment.

Histograms were auto-analyzed using ModFit LT 2.0 in its default configuration. Results were automatically stored in text delimited files.

S-phase error was defined as “calculated minus generated S-phase value”. All data were calculated using Microsoft Excel and plot scales were kept the same for easy comparison of results.

Details of DNA Histogram Generation

RED indicates varied parameter.

1. Number of events

A: **100 - 60,000 events**
Diploid, 80% G0G1 at channel 50, 3% CV
10% S-phase
10% G2M
0% debris
0% aggregates

B: **100 - 60,000 events**
Diploid, 60% G0G1 at channel 50, 3% CV
30% S-phase
10% G2M
0% debris
0% aggregates

2. Number of channels

A: **5 - 100 channels**
Diploid, 80% G0G1, 3% CV
10% S-phase
10% G2M
0% debris
0% aggregates
10,000 events

B: **5 - 100 channels**
Diploid, 60% G0G1 at, 3% CV
30% S-phase
10% G2M
0% debris
0% aggregates
10,000 events

3A. %CV-DNA Diploid

A: **1 - 30% CV**
Diploid, 80% G0G1 at channel 50
10% S-phase
10% G2M
0% debris
0% aggregates
30,000 events

B: **1 - 30% CV**
Diploid, 60% G0G1 at channel 50
30% S-phase
10% G2M
0% debris
0% aggregates
30,000 events

3B. %CV-DNA Aneuploid

A: **1 - 30% CV**
50% Diploid, G0G1at channel 50
50% Aneuploid, G0G1at channel 75
80% G0G1(both)
10% S-phase (both)
10% G2M (both)
0% debris
0% aggregates
30,000 events

B: **1 - 30% CV**
50% Diploid, G0G1at channel 50
50% Aneuploid, G0G1at channel 75
80% G0G1(both)
30% S-phase (both)
10% G2M (both)
0% debris
0% aggregates
30,000 events

4. Aneuploid Fraction

Aneuploid % = 0 - 80
Diploid, G0G1at channel 50
80% G0G1, 3% CV
10% S-phase
10% G2M
Aneuploid, G0G1at channel 75
60% G0G1, 3% CV
30% S-phase
10% G2M
0% debris
0% aggregates
30,000 events

5. %B.A.D.

A:
0 - 90 % debris
0 - 90 % aggregates
Diploid, 80% G0G1 at channel 50, 3 %CV
10 % S-phase
10% G2M
30,000 events

B:
0 - 90 % debris
0 - 90 % aggregates
Diploid, 80% G0G1 at channel 50, 3% CV
30 % S-phase
10% G2M
30,000 events

Details of DNA Histogram Generation

RED indicates varied parameter.

6. DNA Index

A. DNA Index 0.5 to 2.0
50% Diploid
80% G0G1 at channel 50, 3% CV
10% G2M
10% S-phase
0% debris
0% aggregates

50% Aneuploid
80% G0G1, 3% CV
10% G2M
10% S-phase
0% debris
0% aggregates
30,000 events

B. DNA Index 0.5 to 2.0
50% Diploid
80% G0G1 at channel 50, 3% CV
10% G2M
10% S-phase
0% debris
0% aggregates

50% Aneuploid
60% G0G1, 3% CV
10% G2M
30% S-phase
0% debris
0% aggregates
30,000 events

7A. Reduced Chi-Square

Used data generated for 5. %B.A.D.

7B. Reduced Chi-Square

A: 3,000 - 60,000 events
100% Diploid
80% G0G1 at channel 75, 3% CV
10% G2M
10% S-phase
0% debris
0% aggregates

B: 3,000 - 60,000 events
90 % Diploid
80% G0G1 at channel 75, 3% CV
10% G2M (both)
10% S-phase

10% Aneuploid
100% G0G1 at channel 80, 5% CV
0% G2M
0% S-phase
0% debris
0% aggregates

7C. Reduced Chi-Square

Used data generated for 7B.