Sampling

There are times when a reliable measure of current practice is needed (e.g. how often is informed consent being obtained, are care providers conducting post-fall huddles, etc.). One way to obtain this is to take a random sample.

Due to their random nature, samples can never be considered 100% accurate. Two important measures to keep in mind when sampling are the standard margin of error and confidence interval. These are measures of a sample's reliability or accuracy.

The standard margin of error provides a range of accuracy for a given result (e.g. if the sample finds ACP's are being completed 75% of time +/- 5%, the plus/minus number is this statistic's standard margin of error. This means ACP's are likely being completed 70 to 80% of the time.)

Confidence intervals (CI's) deal with the reliability/repeatability of a sample result. If 'several' different samples are taken from the same population, the CI tells us how often these samples would 'match' the *truth* (within the margin of error.) A 95% confidence interval means that if 20 separate random samples were taken, 19 of the 20 samples would match the *true or actual value* (i.e. if all cases were examined).

The larger the sample size, the more reliable the results. Reliability of results will decrease whenever data is split into categories or strata. Political opinion polls are a good example of this. Results from a national opinion poll may find a particular party is expected to win a majority. However, these *same* results are less likely to apply to any one province such as ours (and are even less applicable at an individual riding level).

This is important to keep in mind whenever sample data is presented in categories (e.g. by hospital, patient care unit, gender, age groups, categories of procedure performed, lengths of stay, procedure times, medical specialty, etc.) The more a sample's results are stratified, the less likely it is that the results will represent the *true value*.

As a *rule of thumb*, you should have 30 or more results for any given statistic to feel comfortable with it. Less than 5 results is generally unreliable for interpretation (e.g. if you flip a 'normal' coin 4 times and the same result happens. The odds of having either 4 heads **or** 4 tails is one in eight (12.5%).

The approach used for many statistical tests (e.g. sampling) depend on what is already ***known*** about the topic being examined. Market researchers and academics will sometimes take 'exploratory' samples first – to help decide the specific approach to use later on when the 'formal' test is done. This is often NOT practical in our busy workplace.

If you know the overall size of the population, and do not want to follow a formal statistical protocol, the following may be helpful:

- Assume the population to be randomly sampled is normally distributed (i.e. it is bell-shaped). This is often a reasonable assumption when sample sizes exceed **100**.
- Assume the population variance is unknown.
- If the overall population size is less than 30, the suggested target sample is 100%.
- If the population size exceeds 30, a statistically significant sample size can be calculated.
- The first sample size is usually calculated using a 95% confidence interval and a 5% standard margin of error. If the sample size is too large, you can reduce the 'power' of the sample. In other words, you can decrease the confidence interval and/or increase the margin of error until the sample size required is reasonable given the available resources. A **random number generator, and an Excel spreadsheet are available to calculate the numbers shown on the following page.**

Assuming a ***Standard / Normal Distribution*** with an ***unknown variance.***

Table 1: Calculated Sample Size at 5 % (Standard) Margin of Error

Population Size	Sample Size (at 90, 95 or 99% Confidence Level)		
	90%	**95%**	**99%**
50	29	**33**	38
100	40	**49**	62
150	46	**59**	79
200	50	**65**	91
250	53	**69**	100
500	59	**81**	124
750	62	**85**	136
1000	63	**88**	142
1500	64	**90**	149
2000	65	**92**	153
4000	66	**94**	159

Table 2: Calculated Sample Size at 2.5 % (Standard) Margin of Error

Population Size	Sample Size (at 90, 95 or 99% Confidence Level)		
	90%	**95%**	**99%**
50	42	**44**	46
100	73	**79**	87
150	96	**108**	122
200	115	**132**	154
250	130	**151**	182
500	175	**217**	285
750	198	**254**	352
1000	212	**278**	399
1500	228	**306**	460
2000	237	**322**	498
4000	252	**350**	569

Table 3: Calculated Sample Size at 7.5 % (Standard) Margin of Error

Population Size	Sample Size (at 90, 95 or 99% Confidence Level)		
	90%	**95%**	**99%**
50	19	**23**	30
100	23	**30**	42
150	25	**33**	49
200	26	**35**	54
250	27	**36**	57
500	28	**39**	64
750	29	**40**	67
1000	29	**41**	69
1500	29	**42**	70
2000	29	**42**	71
4000	30	**42**	72