The Ten Principles of Regulatory Compliance Measurement

Richard Fiene PhD, Emeritus Professor of Psychology

Research Institute for Key Indicators/Penn State

March 2023

Abstract

This paper will outline ten principles of regulatory compliance measurement that have been gleaned from 50 years of research into regulatory and licensing databases. For the purposes of this paper, regulatory compliance is to be used interchangeably with licensing and regulatory science. The source of the data is from many jurisdictions in both the United States and Canada. A sampling of these data is displayed on Mendeley Data. These ten principles have been found repeatedly in the various data sets from the jurisdictions that have been analyzed over the past 50 years.

The ten principles to be addressed are the following:

Lack of Variance in data distributions. Data tightly grouped at high compliance levels.

Ceiling/Plateau Effect in data distributions.

Difficulty distinguishing levels of quality between full and substantial compliance.

Nominal measurement level: Either In-Compliance or Out-of-Compliance.

Attempting to move to ordinal measurement level when quality is included.

Dichotomization of data is warranted because of the data distribution.

Problem with false negatives and positives, especially false negatives.

Lack of reliability and validity testing.

Ease in distinguishing levels of quality between low and substantial compliance.

Skewed Data. Majority of programs in substantial or full regulatory compliance.

The first principle deals with the lack of Variance in data distributions. Data are found to be tightly grouped at high compliance levels (upper 90% level). This will lead to another principle addressed later in this paper dealing with skewness of the data distribution. In fact, the majority of scores are at a full regulatory compliance level, in other words, 100% in compliance with all rules and regulations. This led to variance statistics showing little movement and the majority of programs being in very close proximity. This makes for difficult statistical analyses when there is little variance in the data set.

The second principle is finding a ceiling or plateau effect in data distributions. It was like there was a diminishing returns effect as one moves from substantial regulatory compliance (upper 90%+) to full regulatory compliance (100%) with all rules and regulations. This was especially true when one compares the regulatory compliance levels with program quality scores on those same programs which is addressed more in the next principle.

The third principle is the difficulty distinguishing levels of quality between full and substantial compliance. This principle builds off of the previous principle dealing with a ceiling or plateau effect. Because so much of the data, as much as 70-80% of programs, are grouped so tightly at the substantial and high levels of regulatory compliance when one begins to go beyond regulatory compliance and begin to look at quality there is a great deal of difficulty distinguishing levels of quality. In other words, the full regulatory compliant level programs are not necessarily the highest quality programs.

The fourth principle is the fact that rules and regulations are measured at a nominal measurement level: the rules and regulations are either In-Compliance or Out-of-Compliance. The rule or regulation is measured at a "Yes" or "No" level or a "1" or "0" level. There are no in-between measures, no ordinal measurement going on. Either you got it, or you don't. It is black or white, no shades of gray. It is just the nature of measurement when it comes to rules and regulations which are very different in other measurement systems. The data are very discrete and not continuous. They are frequency counts and not a ruler type of measurement. One will not find an interval level of measurement in any regulatory science data distribution.

A fifth principle is attempting to move to an ordinal measurement level when quality is included. This principle builds off of the previous principle in which in some cases it has been suggested to add a quality component to particular rules or regulations. This is an interesting development and moves the philosophy from one of "Do no harm" to one of "Do things well". It will be interesting to see how much this concept moves forward and changes a basic tenet in the regulatory science field which is more based upon health & safety, gatekeeper, hard data, risk aversion, and deficit based.

The sixth principle of regulatory compliance measurement is the ability to dichotomize the data can be warranted because of the data distribution. Data dichotomization is generally not recommended because it accentuates differences in a data set. However, given the nature of regulatory compliance measurement being at a nominal level, fitting into a bucket format, the lack of variance, and the skewness of the data distribution all lead to the ability to dichotomization of the data set.

The seventh principle has to do with the problem with false negatives and positives, especially false negatives. Because of the data being measured in a nominal In-Compliance vs Out-of-Compliance dichotomy it can lead to false negatives in which In-Compliance decisions are made that in reality are not In-Compliance. False positives are a problem as well but not as much of a problem as false negatives. In false positives, Out-of-Compliance may be determined when in reality the rule or regulation is actually In-Compliance. This is not a good scenario for the provider of services, but it potentially doesn't harm the client as much as when a false negative occurs.

The eighth principle is the lack of reliability and validity testing. This principle builds from the previous principle in that there are very few examples of scientific testing of instrumentation and the administration of protocols to make certain that everything is running as it should. Because of this, it leads to the above problem of false positives and negatives. All jurisdictions need to build in regular reliability and validity testing to ascertain that the final decision making is within the ranges that are acceptable.

The ninth principle is the ease in distinguishing levels of quality between low and substantial compliance. The one result that has been consistent over the years is the ability to see differences in programs that score low on regulatory compliance versus those that are at a substantial or high compliant level. From a licensing or regulatory administration point of view this is a real plus in being able to be an effective gatekeeper and keeping non-optimal programs out of service. But as indicated in the third principle this advantage is short-lived as one moves up the regulatory compliance scale to substantial and finally to full regulatory compliance. When one gets to these levels it becomes increasingly difficult to distinguish differences in quality in those programs that are in substantial regulatory compliance versus those that are in full regulatory compliance. It appears that the regulatory compliance theory of diminishing returns is rearing its plateau/ceiling effect. The policy implications are immense since the assumption is that there is a linear relationship between program quality and regulatory compliance. How do we more effectively deal with this non-linear relationship in formulating public policy regarding licensing decision making?

And the final tenth principle is that regulatory compliance data are always skewed data. The majority of programs are in substantial or full regulatory compliance. And in many cases, this can be rather severe. There generally is a long tail which contains some low regulatory compliant programs, but these are usually few in number. The data distribution just does not approach a normally distributed curve as we see in many other examples of social science data distributions.

It is important as the regulatory science field moves forward that we remain cognizant of the limitations of regulatory compliance measurement. There are some severe limitations that need to be addressed (e.g., skewed data, lack of variance in data, ceiling effect, nominal metrics) and mitigated (e.g., data dichotomization) or it will continue to lead to problems in our analyses (e.g., false positives and negatives).

Richard Fiene PhD is a research psychologist and regulatory scientist at the Research Institute for Key Indicators. He is also an Emeritus Professor of Psychology at the Penn State Edna Bennett Pierce Prevention Research Center and a senior research consultant at the National Association for Regulatory Administration. Contact Dr Fiene at rfiene@rikinstitute.com