

# Hawaii QRIS Key Indicator Blueprint Report

**Richard Fiene, Ph.D.**

**August 30, 2013**

## **ABSTRACT**

This report will provide a blueprint for Hawaii's QRIS in developing a key indicator approach to help streamline their present assessment process. The report will be organized into the following major headings: an introduction to the key indicator methodology; how key indicators fit into the larger program monitoring of early care and education programs; how key indicators will be applied to QRIS and to Hawaii's QRIS in particular; the technical aspects of the key indicator methodology, the sample to be drawn from the population, although the full population of early care and education programs may be able to be used; potential results from the analyses; a timeline for this developmental effort; and potential cost savings from the approach. This blueprint report will answer all the questions about developing key indicators for QRIS, the what, how, why, when, etc...

## **INTRODUCTION**

The Key Indicator Methodology was developed to help streamline the program monitoring of early care and education programs. It was first applied in child care licensing (Fiene & Nixon, 1985) but has been used in many other service types, such as: Head Start Performance Standards (Fiene, 2013a), National Accreditation (Fiene, 1996), and child and adult residential programs (Kroh & Melusky, 2010). The methodology is based upon statistical protocols that have been developed in the tests and measurements literature in which an abbreviated set of items is used to statistically predict as if the full test was applied. This methodology has been used in regulatory analysis and more recently has been proposed for use in Quality Rating and Improvement Systems (QRIS) (Fiene, 2013b).

## DIFFERENTIAL PROGRAM MONITORING

Key indicators are an important component of differential program monitoring which employs an abbreviated review rather than a comprehensive or full review of a program. It is one of several key elements that have been identified in the research literature to help improve the cost effectiveness and efficiency of the program monitoring of early care and education programs (Fiene, 2013b, c)(See the Appendix). A recent addition to differential monitoring are QRIS – Quality Rating and Improvement Systems. Key indicators have a long history of development within the licensing literature (Fiene & Kroh, 2000) but have only recently been proposed to be used with QRIS. This proposed blueprint is a first for a state to determine the feasibility of using the key indicator approach with its QRIS system.

The other key elements of the differential program monitoring approach are the following: program compliance/licensing which is generally a state's health and safety rules/regulations that govern child care. At the national level this would be *Caring for Our Children: National Performance Standards for Health and Safety in Child Care (2012)*. The program quality key element is generally represented by the state's QRIS. At the national level it is represented by accreditation, such as NAEYC, NECPA, or NAFCC. The key indicator element is represented by the state's statistical predictor rules/regulations drawn from their comprehensive set of health and safety rules/regulations that govern child care. At the national level, an example is the *13 Indicator of Quality Child Care (2002)*. This element can also represent a state's statistical predictor QRIS standards drawn from the comprehensive set of QRIS standards. The purpose of this Blueprint Report is to develop these statistically predictor QRIS standards. The last key element to be addressed in this report is the risk assessment key element in which these are the high risk rules/regulations that place children at greatest risk of mortality or morbidity. At the national level, an example is *Stepping Stones to Caring for Our Children (2013)*. These are generally determined via a weighting system in licensing or a point system with QRIS.

## KEY INDICATORS APPLIED TO HAWAII'S QRIS

Hawaii's QRIS is somewhat unique in that its assessment system is drawn very heavily from off-the-shelf assessment tools, such as the ERS's, CLASS, PAS/BAS in addition to QRIS program standards. This will pose significant challenges because of the psychometric properties of these standardized tools. However, with that said, the key indicator methodology is drawn directly from the tests and measurements research literature in which it is an approach in taking a comprehensive test and reducing it down to a group of statistical predictor items. The key indicator methodology will not alter the scale structure of any of the assessment tools. The purpose of the key indicator methodology is to establish a protocol

so that a determination of a full score and the appropriate level can be statistically predicted from a smaller set of items from that respective tool, in Hawaii's QRIS standards, ERS's, CLASS, PAS/BAS, NAEYC, NAFCC.

The key indicators can eventually be tied to the professional development/training/technical assistance system to link resources to specific needs of the programs. It also has the capability of tying them to an early learning benchmarking and child outcomes at some point in the future. This would be accomplished in the full implementation of the Differential Monitoring Logic Model and Algorithm (DMLMA©) as depicted in the Appendix.

### TECHNICAL ASPECTS OF THE KEY INDICATOR METHODOLOGY

This section provides the technical and statistical aspects of the key indicator methodology. It will provide the roadmap in taking the Hawaii QRIS data base through the necessary steps to generating the respective key indicators.

One of the first steps is to sort the data into high and low groups, generally the highest and lowest ratings can be used for this sorting. In very large states this is done on a sampling basis but in Hawaii's case we should be able to use all the programs who participate in the QRIS and not take a sample. Frequency data will be obtained on those programs in the top level (usually top 20-25%) and the bottom level (usually the bottom 20-25%). The middle levels are not used for the purposes of these analyses. These two groups (top level & the bottom level) are then compared to how each program scored on each item within the specific assessment tool (see Figure 1). An example would be the following: Item 16 from the ECERS – Encouraging Children to Communicate. Sort all the providers by the number in the highest group and the lowest. Then determine how each program scored on item 16, did they get a 5 or higher or a 3 and lower? Fill in the cells within Figure 1 accordingly (see Figure 2).

<b>Figure 1</b>	<i>Providers In Compliance or Top 25%</i>	<i>Programs Out Of Compliance or Bottom 25%</i>	<i>Row Total</i>
<i>Highest level (top 20-25%)</i>	<i>A</i>	<i>B</i>	<i>Y</i>
<i>Lowest level (bottom 20-25%)</i>	<i>C</i>	<i>D</i>	<i>Z</i>
<i>Column Total</i>	<i>W</i>	<i>X</i>	<i>Grand Total</i>

Figure 2 depicts that all programs that were in the top 25% (5+ on ECERS, Item 16) were also in the highest rating while the bottom 25% (3 or lower on the ECERS, Item 16) were also in the lowest rating. The data depicted in Figure 2 are taken from studies completed in Pennsylvania in 2002 (Fiene, etal) and 2006 (Barnard, Smith, Fiene & Swanson) in which their quality rating and improvement system (QRIS), Keystone STARS, was validated.

<b>Figure 2 – Pa. Study (Fiene, etal, 2002).</b>	<b><i>Providers In Compliance or Top 25%</i></b>	<b><i>Programs Out Of Compliance or Bottom 25%</i></b>	<b><i>Row Total</i></b>
<b><i>Highest Star level in Pa.</i></b>	<b><i>117</i></b>	<b><i>0</i></b>	<b><i>117</i></b>
<b><i>Lowest Star level in Pa.</i></b>	<b><i>0</i></b>	<b><i>35</i></b>	<b><i>35</i></b>
<b><i>Column Total</i></b>	<b><i>117</i></b>	<b><i>35</i></b>	<b><i>152</i></b>

Once the data are sorted in the above matrix, the following formula (Figure 3) is used to determine if Item 16 is a key indicator or not by calculating its respective Phi coefficient. Please refer back to Figure 1 for the actual placement within the cells and Figure 2 for the data within the cells. The legend (Figure 4) below the formula shows how the cells are defined.

**Figure 3 – Formula for Phi Coefficient**

$$\phi = (A)(D) - (B)(C) \div \sqrt{(W)(X)(Y)(Z)}$$

**Figure 4 – Legend for the Cells within the Phi Coefficient**

- A = High Group + Programs in Compliance on Specific Compliance Measure.***
- B = High Group + Programs out of Compliance on Specific Compliance Measure.***
- C = Low Group + Programs in Compliance on Specific Compliance Measure.***
- D = Low Group + Programs out of Compliance on Specific Compliance Measure.***
  
- W = Total Number of Programs in Compliance on Specific Compliance Measure.***
- X = Total Number of Programs out of Compliance on Specific Compliance Measure.***
- Y = Total Number of Programs in High Group.***
- Z = Total Number of Programs in Low Group.***

Once the data are run through the formula in Figure 3, the following chart (Figure 5) can be used to make the final determination of including or not including the item as a key indicator. Based upon the chart in Figure 5, it is best to have a Phi Coefficient approaching +1.00 since we are dealing with normally distributed data<sup>1</sup>. This requirement is relaxed with licensing rules & QRIS selected standards only (+.26 and higher) because the data are more skewed but this should not be the case as much with Hawaii's Quality Rating and Improvement System (QRIS) data because the measures selected in the QRIS are mostly standardized tools with more normally distributed data.

Continuing with the chart in Figure 5, if the Phi Coefficient is between +.25 and -.25, this indicates that the indicator is unpredictable in being able to predict overall compliance with the quality rating assessment tool. Either a false positive in which the indicator appears too often in the low group as being in compliance, or a false negative in which the indicator appears too often in the high group as being out of compliance<sup>2</sup>. This can occur with Phi Coefficients above +.25 but it becomes unlikely as we approach +1.00 although there is always the possibility that other standards/rules/regulations could be found out of compliance (this was demonstrated in a study conducted by the author (Fiene, 2013c) with Head Start programs). Another solution is to increase the number of key indicators to be reviewed but this will cut down on the efficiency which is desirable and the purpose of the key indicators.

The last possible outcome with the Phi Coefficient is if it is between -.26 and -1.00, this indicates that the indicator is a terrible predictor because it is doing just the opposite of the decision we want to make. The indicator would predominantly be in compliance with the low group rather than the high group so it would be statistically predicting overall non-compliance. This is obviously something we do not want to occur.

**Figure 5 – Thresholds for the Phi Coefficient (Fiene & Nixon, 1983, 1985)**

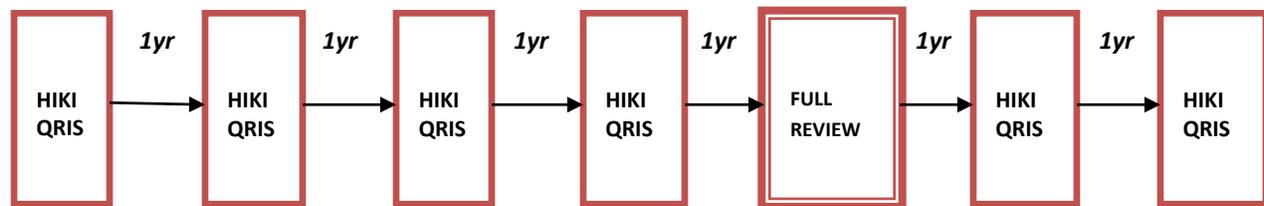
<u>Phi Coefficient Range</u>	<u>Characteristic of Indicator</u>	<u>Decision</u>
<b>(+1.00) – (+.26)</b>	<b>Good Predictor</b>	<b>Include</b>
<b>(+.25) – (-.25)</b>	<b>Unpredictable</b>	<b>Do not Include</b>
<b>(-.26) – (-1.00)</b>	<b>Terrible Predictor</b>	<b>Do not Include</b>

The key indicators should then only be used with those programs who have attained the highest rating. It is not intended for those programs that have attained lower ratings. However, even with those programs that have attained the highest rating, every 3-5 years a full, comprehensive

review using the full assessment tools and QRIS standards should occur (see Figure 6 for a graphical depiction). It is intended that a re-validation of the key indicators occur on a periodic basis to make certain that the key indicators have not changed because of differences in compliance history. This is an important and necessary step for the state to engage in to ascertain the overall validity and reliability of the assessment system. Also there should not have been any major changes in the program while the key indicators are being administered, such as the director leaving or a large percentage of teachers leaving or enrollment increasing significantly, or a change in the licensing status of the program.

**Figure 6 - Proposed DMLMA System with Key Indicators (KI)**

*Use of Hawaii Key Indicators (HIKI) for QRIS with a Full Review every 4<sup>th</sup> Year*



**SAMPLE**

Generally a sample is drawn from the population of early care and education facilities in the respective state. With this being said, the chances are the full population will be able to be used in Hawaii’s case because of the manageable number of facilities. This should be able to be done with centers as well as with homes<sup>2a</sup>.

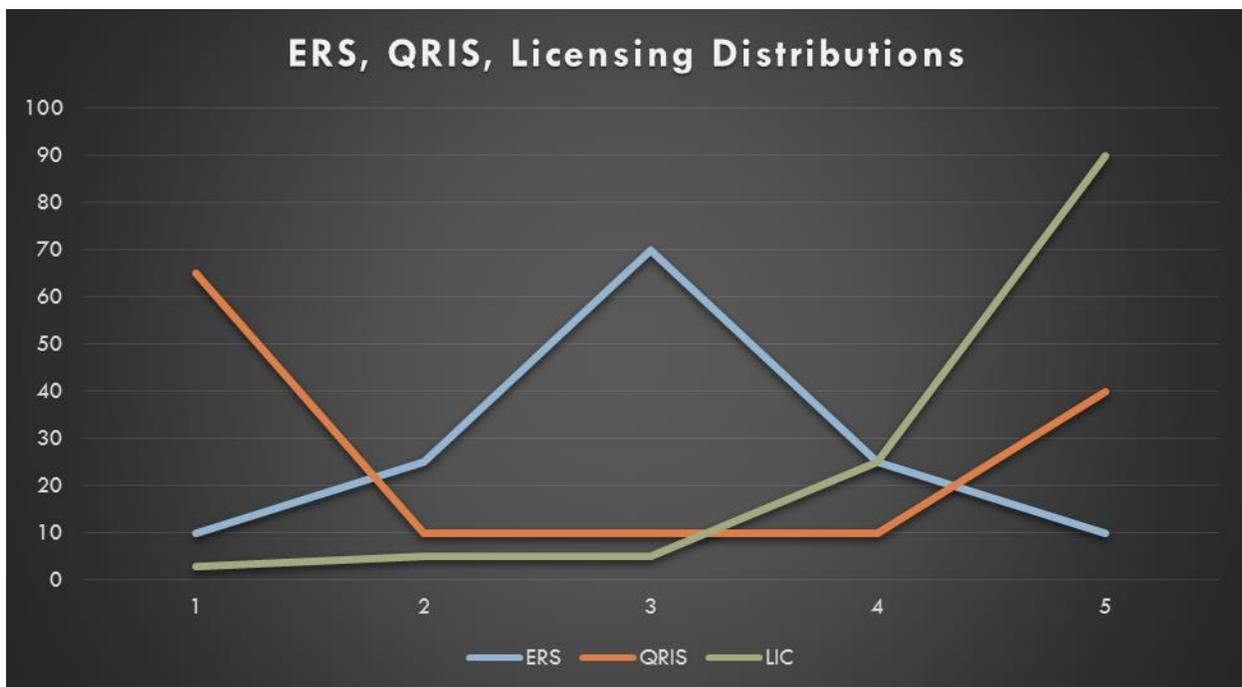
**POTENTIAL RESULTS**

The potential results are drawn from previous studies conducted by the author (Fiene, 2013b) in which key indicators were generated for the ECERS-R and FCCERS-R. All the specific items in the ECERS-R and FCCERS-R were run through the Phi Coefficient formula in Figure 3 above after having sorted the data into a high group (5 or higher) and a low group (3 or less) for the overall ECERS-R and FCCERS-R scores. This same procedure will be followed with the Hawaii QRIS but in this case the individual ERS item score will be compared with the respective Star Levels which will be sorted into a high group (top Level) and a low group (bottom Level) in order to determine which individual ERS items become key indicators. This process will be repeated for all ERS items and then extended to CLASS and PAS/BAS items as well as QRIS standards and where appropriate to NAEYC and NAFCC items.

It is estimated from previous studies (Fiene, 2013a; 2013c; 2013d) that approximately 10% of the ERS, CLASS, PAS/BAS, NAEYC, NAFCC items & QRIS standards will become key indicators. If this holds true it will substantially reduce the total number of items to review for QRIS assessments. It is also expected that the Phi Coefficients will be very high at a .90 level or higher because of the dichotomization of the data which should be normally distributed rather than significantly skewed. Also there will be significant redundancy in the data because the rating levels are so much tied to the standardized assessments in that the ERS, CLASS, PAS/BAS, NAEYC, and NAFCC are directly cross-walked to increasing rating levels.

As mentioned earlier, the measurement issues with the various standardized tools will provide challenges because of their data distributions. In the past when key indicators have been generated with licensing data which are highly skewed, dichotomization of the data is regularly done. However, when one looks at Figure 7 it is clear that the standardized assessments are more normally distributed than skewed<sup>3</sup>. Generally dichotomization of data should not be done with normally distributed data<sup>4</sup>; however, in this case with Hawaii’s QRIS and how the standardized assessments are used to make decisions regarding rating levels, it is appropriate to do so since the data lend themselves to being sorted into discrete categories, such as rating levels. The dichotomization will compare the lowest rating level with the highest rating level in order to generate the key indicators.

**Figure 7 – Data Distribution Comparisons of ERS, QRIS, and Licensing Data**



**TIMELINE**

As soon as all early care and education programs have gone through their assessment phase, it will be possible to do the calculations to determine the Phi Coefficients and generate the key indicators. I am guessing that this should not take any longer than 1 year but could be completed in a much shorter period of time if the assessments on individual programs could be moved up (see Figure 8). The analytical phase should take no longer than a month with an additional month to write up the report. A face to face presentation of the analyses could be done after these two months.

The timeline presented in Figure 8 can be adjusted to the specific needs of Hawaii’s QRIS system. The timeline is based upon previous projects and the average time to generate key indicators. Another consideration or task is the development of the policies and procedures to be developed and implemented regarding the use of key indicators. This was not specifically listed on the timeline because it is something that is generally developed throughout the project with feedback from all the stakeholders who will be impacted by the use of this new approach to assessment and monitoring.

**Figure 8 - HAWAII QRIS KEY INDICATOR (KI) PROJECT TIMELINE**

<u>TASK</u>	<u>MONTHS</u>						
	1	2	3	4	5	6	7
<b>Collect Data</b>	XXXXXXX						
<b>Sort Data</b>		XXXXXXX					
<b>Run Analyses</b>			XXXXXXX				
<b>Generate KI</b>				XXXXXXX			
<b>Training on KI</b>					XXXXXXX		
<b>KI Reliability</b>						XXXXXXX	
<b>Implementation</b>							XXXXXXX

**Legend:**

**Collect Data** – dependent upon the total number of programs participating it would be determined to collect data on all participants or to complete a sample.

**Sort Data** – the individual programs are sorted into high and low groups representing the top 25% and the bottom 25% of programs as they have scored on the respective assessment tools and standards.

**Run Analyses** – each individual item within each of the assessment tools for every program will be compared to the sorting process of the high and low groups.

**Generate KI** – a 2 x 2 matrix is constructed and the key indicators (KI) are generated from this matrix through the use of a phi coefficient. A final report will be delivered to Hawaii executive staff.

**Training on KI** – all staff who will be using the KI will be trained on its use.

**KI Reliability** – reliability will be established by having two staff go out together and administer the key indicators separately and comparing their results.

**Implementation** – once reliability has been established, full implementation will begin.

## COST SAVINGS

Again based upon previous studies most recently completed in California in 2010

(<http://www.mycccl.ca.gov/res/docs/12022010HandoutStakeholderMeeting.pdf>), time savings of 50% have been attained by using a key indicator or abbreviated tool in completing assessments. It only makes sense that if an assessment can be completed in one hour rather than 2 – 4 hours that a state will see time savings. It is being assumed that equivalent savings should also be the case with Hawaii's QRIS although this cannot be made certain until the new key indicator or abbreviated tool is actually used for a period of time. Once the new key indicators are used for several months, comparisons could be made to when the full assessments were done.

## CONCLUSION AND NEXT STEPS

This blueprint report has given the basic parameters to develop a key indicator approach to Hawaii's QRIS assessment tools. By following this blueprint Hawaii staff should be able to fully implement the approach. Hawaii staff would also need to determine if they have the internal capability for the development of the key indicators or if there will be the need to outsource certain aspects of the development. This will be an important consideration as Hawaii moves forward with this project. I have provided two options for your consideration in moving forward.

### *Option 1 – Development of System Internally:*

This would require either information systems or research & evaluation staff to analyze the data, generate key indicators for each assessment tool, and training of staff. I could provide the necessary consulting services to help the staff work through the methodology. This would probably require at least one face to face meeting with regular monthly conference calls between myself and staff. Discussions of the formatting of data and the types of analyses would be discussed and demonstrated.

### *Option 2 – Development of System Externally:*

In this option I could do all the methodological work demonstrating how I would need the data sent to me, the analytical work in generating key indicators for each assessment tool, a report

detailing the methodology and results. The only thing that Hawaii staff would need to do is get the data to me, all other aspects of what is delineated in the timeline in Figure 8 would be completed by me. This would probably require several face to face trips to explain the process, the results, and do training of staff. Once everything was in place, Hawaii staff would have a fully implemented system.

If the above options are of interest I can provide detailed budgets for either one or both.

Notes:

- 1, 4. The reason for pointing out the need to have a higher Phi Coefficient than what has been reported previously (Fiene & Nixon, 1983, 1985) is the fact that the dichotomization of data should only be used with skewed data and not normally distributed data because it will accentuate differences. However, since the purpose of the dichotomization of data is only for sorting into a high and low group, it would appear to be acceptable for this purpose (MacCallun, etal, 2002. On the practice of dichotomization of quantitative variables, *Psychological Methods*, 7, 1, 19-40.).
2. These results would show an increase in cells B and C in Figure 1 which is undesirable; it should always be the case where  $A + D > B + C$  for key indicators to maintain their predictive validity.
  - 2a. If a sample must be drawn, I can help to provide the guidance in pulling such a sample.
3. The distinction between making decisions with skewed (Licensing) as versus normally distributed (ERS) data is an important one because there is a greater likelihood with skewed data of introducing less than optimal programs into the high group when sorting programmatic data into high and low groups. This then makes it more difficult to identify the best programs. However, because of the distribution with skewed data the same cannot be said with the low group in which case it is relatively easy to identify the problem programs. This is not as much of a concern when the data are more normally distributed in which it is relatively easy to identify both the optimal and problem programs. This is an excellent example of the need of weighting of standards in order to increase the normal distribution of the data.

**REFERENCES AND ADDITIONAL RELATED READINGS REGARDING DMLMA:**

- Barnard, Smith, Fiene, Swanson (2006). Evaluation of Pennsylvania's Keystone STARS Quality Rating and Improvement System, Pittsburgh: Pennsylvania, Office of Child Development.
- Fiene (2013a). A comparison of international child care and US child care using the Child Care Aware – NACCRRA (National Association of Child Care Resource and Referral Agencies) child care benchmarks, *International Journal of Child Care and Education Policy*, 7(1), 1-15.
- Fiene (2013b). *Differential monitoring logic model and algorithm*. Middletown: Pennsylvania, Research Institute for Key Indicators.
- Fiene (2013c). Head Start Key Indicators. Middletown: Pennsylvania, Research Institute for Key Indicators.
- Fiene (2013d). Kansas Child Care Key Indicators. Middletown: Pennsylvania, Research Institute for Key Indicators.
- Fiene (2013e). Validation of Georgia's core rule differential monitoring system. Middletown: Pennsylvania, Research Institute for Key Indicators.
- Fiene (2007). Child Development Program Evaluation & Caregiver Observation Scale, in T Halle (Ed.), *Early Care and Education Quality Measures Compendium*, Washington, D.C.: Child Trends.
- Fiene (2003). Licensing related indicators of quality child care, *Child Care Bulletin*, winter 2002-2003, pps 12-13.
- Fiene (2002a). *Thirteen indicators of quality child care: Research update*. Washington, DC: Office of the Assistant Secretary for Planning and Evaluation, US Department of Health and Human Services.
- Fiene (2002b). Improving child care quality through an infant caregiver mentoring project, *Child and Youth Care Forum*, 31(2), 75-83.
- Fiene, Iutovich, Johnson, & Koppel (1998). Child day care quality linked to opportunities for professional development: An applied community psychology example. *Community Psychologist*, 31(1), 10-11.
- Fiene (1996). Using a statistical-indicator methodology for accreditation, in *NAEYC Accreditation: A Decade of Learning and the Years Ahead*, S. Bredekamp & B. Willer, editors, Washington, D.C.: National Association for the Education of Young Children.
- Fiene (1995). Utilizing a statewide training system to improve child day care quality: The other system in a program quality improvement model. *Child Welfare*, Volume LXXIV, #6, November-December, 1189-1201.

- ❑ Fiene (1985). Measuring the effectiveness of regulations, *New England Journal of Human Services*, 5(2), 38-39.
- ❑ Fiene (1981). A new tool for day care monitoring introduced by children's consortium, *Evaluation Practice*, 1(2), 10-11.
- ❑ Fiene, Greenberg, Bergsten, Carl, Fegley, & Gibbons (2002). *The Pennsylvania early childhood quality settings study*, Harrisburg, Pennsylvania: Governor's Task Force on Early Care and Education.
- ❑ Fiene & Kroh (2000). Licensing Measurement and Systems, *NARA Licensing Curriculum*. Washington, D.C.: National Association for Regulatory Administration.
- ❑ Fiene & Nixon (1985). Instrument based program monitoring and the indicator checklist for child care, *Child Care Quarterly*, 14(3), 198-214.
- ❑ Griffin & Fiene (1995). *A systematic approach to policy planning and quality improvement for child care: A technical manual for state administrators*. Washington, D.C.: National Center for Clinical Infant Programs-Zero to Three.
- ❑ Kontos & Fiene (1987). Child care quality, compliance with regulations, and children's development: The Pennsylvania Study, in *Quality in Child Care: What Does Research Tell Us?* Phillips, editor, Washington, D.C.: National Association for the Education of Young Children.
- ❑ Zellman, G. L. and Fiene, R. (2012). *Validation of Quality Rating and Improvement Systems for Early Care and Education and School-Age Care*, Research-to-Policy, Research-to-Practice Brief OPRE 2012. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services

**For additional information regarding this report, please contact:**

Richard Fiene, Ph.D., Director/President  
Research Institute for Key Indicators (RIKI)  
41 Grandview Drive  
Middletown, PA. 17057  
[DrFiene@gmail.com](mailto:DrFiene@gmail.com)  
[RIKI.Institute@gmail.com](mailto:RIKI.Institute@gmail.com)  
717-944-5868 Phone and Fax  
<http://RIKInstitute.wikispaces.com>  
<http://pennstate.academia.edu/RickFiene>

### Appendix

#### DIFFERENTIAL MONITORING LOGIC MODEL & ALGORITHM (DMLMA©) (Fiene, 2012): A 4<sup>th</sup> Generation ECPQIM – Early Childhood Program Quality Indicator Model

$$CI \times PQ \Rightarrow RA + KI \Rightarrow DM + PD \Rightarrow CO$$

**Definitions of Key Elements:**

- PC = Program Compliance/Licensing (Health and Safety) (*Caring for Our Children*)
- PQ = QRIS/Accreditation/Caregiver/Child Interactions/Classroom Environment Quality (*ERS/CLASS/PAS/BAS*)
- RA = Risk Assessment, (High Risk Rules) (*Stepping Stones*)
- KI = Key Indicators (Predictor Rules) (*13 Key Indicators of Quality Child Care*)
- DM = Differential Monitoring (How often to visit and what to review)
- PD = Professional Development/Technical Assistance/Training (Not pictured but part of Model)
- CO = Child Outcomes (Not pictured but part of Model)

