A Guide to Interpreting Hot Spot Reports

The Center for State Child Welfare Data
Rev. December 2015

Table of Contents

ABOUT THIS DOCUMENT

This manual provides guidance on how to use the Hot Spot function of The Center for State Child Welfare Data's Multistate Foster Care Data Archive (FCDA) web tool (https://fcda.chapinhall.org/). The document is organized into the following sections:

Brief summary. This overview offers a concise, general explanation of the information that Hot Spot reports provide and why that information is relevant to child welfare agencies' decisions about how to allocate resources toward outcome improvement efforts.

Background. This section goes into further detail regarding the purpose of the Hot Spot program and how its results can be applied to the Continuous Quality Improvement (CQI) process. This section also includes a brief primer on risk-adjustment—the statistical method used in Hot Spot calculations.

Selecting a report & Interpreting the results. These sections contain instructions for how to run and interpret Hot Spot reports. We are grateful to the State of Tennessee for giving us permission to use their data for educational purposes.

Appendix A: Hot Spot Methodology. This appendix contains the analytic specifications for Hot Spot calculations.

Appendix B: Evidence use throughout the CQI process. This appendix contains general information on the role of evidence throughout the CQI process and resources for further education.

BRIEF SUMMARY

Hot Spot is an analytic tool available to members of The Center for State Child Welfare Data via the Multistate Foster Care Data Archive (FCDA) web tool (https://fcda.chapinhall.org). For each member state, Hot Spot uses a **risk-adjustment** method to calculate the extent to which a county's performance on a particular outcome differs from the state and national performance *after* controlling for influential child- and county-level factors. The outcomes currently available for analysis on Hot Spot are:

- Foster care placement rate per 1,000 children
- Percent of children exiting to permanency within 12, 24, 36, and 48 months (outcomes may be further stratified by child age)

Risk-adjustment provides a fair way of comparing jurisdictions to one another by recognizing two important realities. The first is that within a given state, a particular child welfare outcome will vary from county to county. The second is that child welfare outcomes depend on more than simply the county where a child lives, and those other factors also vary from place to place.

In an effort to "level the playing field," risk-adjustment asks the question: If I remove the effect of these other factors on the outcome of interest, will there still be differences in county performance on that outcome? If there are still significant differences between counties after accounting for those other influences, we can say that performance on the outcome is, in fact, associated with the specific counties in which children are placed.

That analysis is important because it provides child welfare administrators with knowledge about which populations of children within the state have better and worse than average outcomes, all other things being equal. Such information is invaluable as states aim to target performance improvement resources toward groups of children who stand most to benefit.

BACKGROUND

What is Hot Spot and what information does it provide?

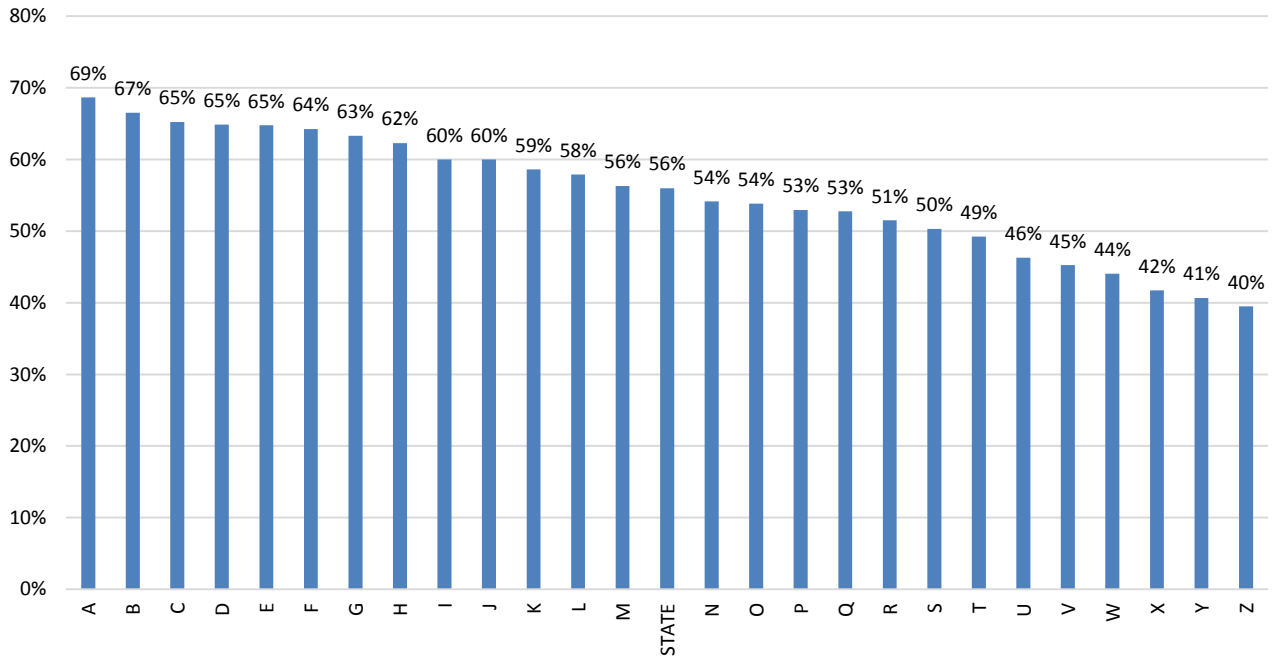
Hot Spot is an analytic tool available to members of The Center for State Child Welfare Data via the FCDA web tool (https://fcda.chapinhall.org). For each member state, Hot Spot uses a **risk-adjustment** method to calculate the extent to which a county's performance on a particular outcome differs from the state and national performance *after* controlling for influential child- and county-level factors.

What is risk-adjustment and why do it?

A quick overview of how risk-adjustment works sheds light on what Hot Spot results mean and how they can be applied. The method provides a fair way of comparing jurisdictions to one another by recognizing two important realities.

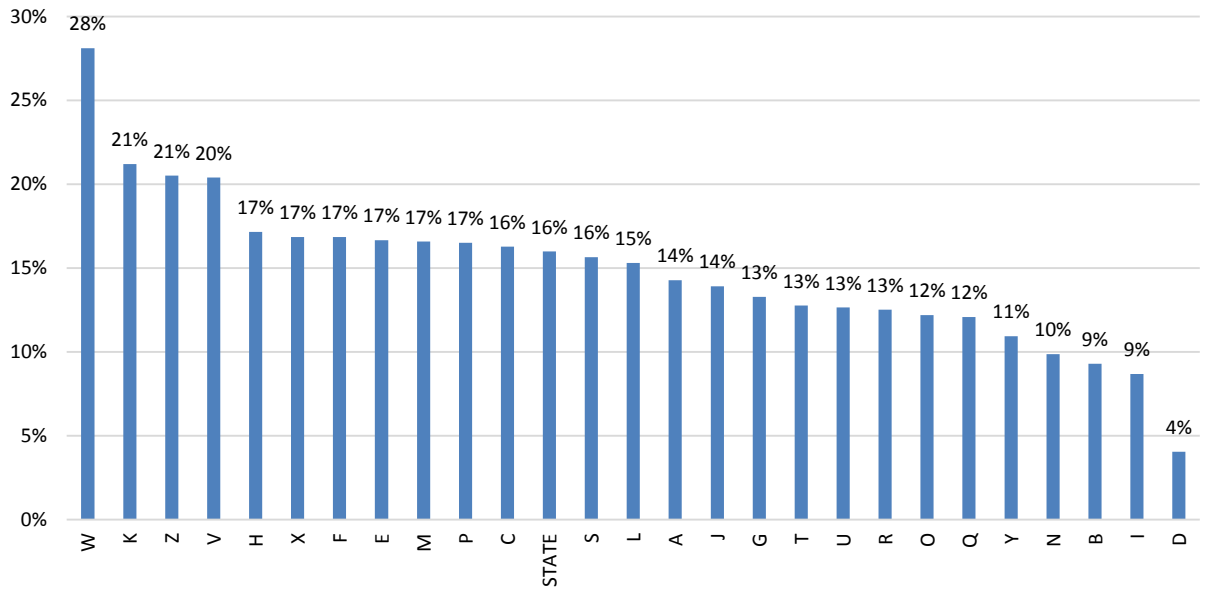
The first reality is that, **within a given state, a particular child welfare outcome will vary from county to county.** Consider Figure 1 below, which shows the percent of children exiting to permanency within 12 months for 26 counties inside a sample state. Some counties have a higher permanency rate than others. Specifically, in this state, the likelihood of exiting to permanency within 12 months ranges from 69% in County A to 40% in County Z. The statewide rate of permanency within 12 months is 56%.

Figure 1: Percent of children exiting to permanency within 12 months, by county (sample state)



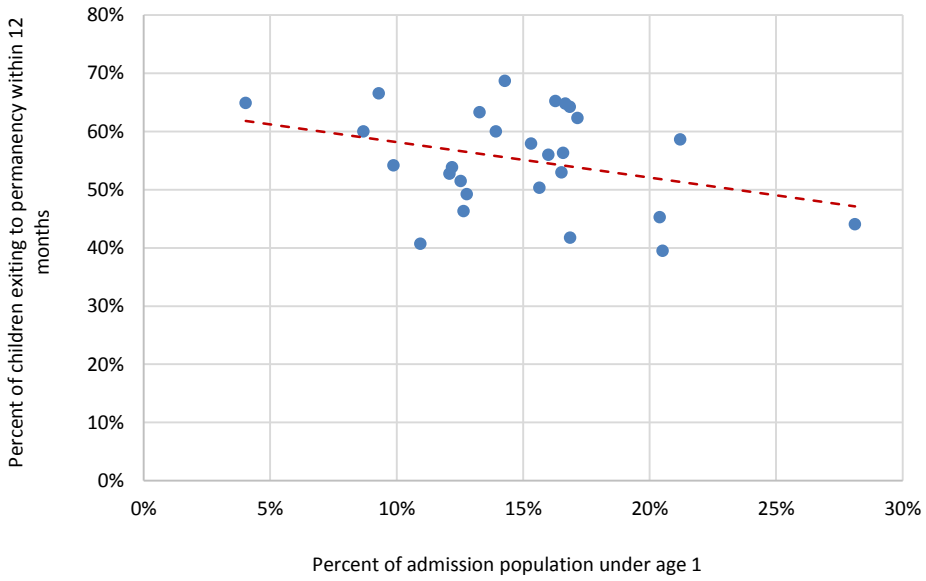
The second reality is that **child welfare outcomes depend on more than simply where a child lives, and those other factors also vary from place to place.** Consider the child's age at entry, for example. Some counties' admission populations have a large proportion of infants; some counties have proportionally fewer infants. Figure 2 shows the proportion of infants in the admission population for each of the 26 counties noted above. As expected, the proportional representation of infants varies. In County W, 28% of children entering foster care are under the age of one; in County D, only 4% of children entering care are that age. Statewide, 16% of children entering care are infants.

Figure 2: Percent of admissions population under the age of 1, by county (sample state)

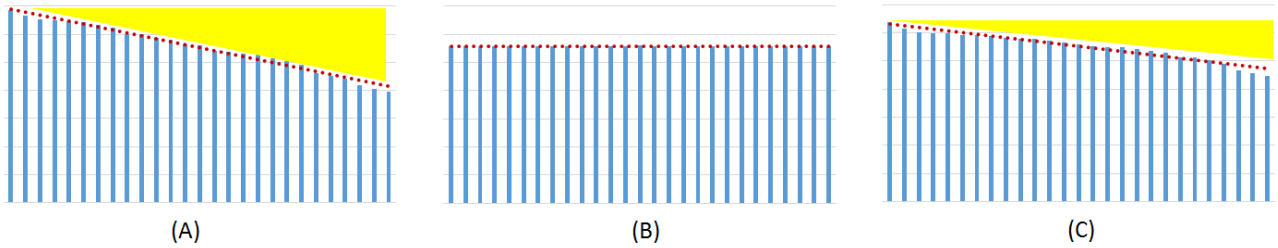


Research has consistently shown that the likelihood of achieving permanency within 12 months is heavily influenced by a child's age at entry to foster care; specifically, children who enter as infants are typically less likely to achieve permanency within one year than older children are (in large part because they are more likely than older children to exit to adoption—a permanency solution that takes longer than, say, reunification). Therefore, a county where infants make up a large proportion of the admissions population will probably have a lower rate of permanency within 12 months than a county with proportionally fewer infants. A rough look at these figures side by side in the sample state shows this to be true: Generally speaking, as the county's concentration of infants goes up, its permanency rate goes down (Figure 3).

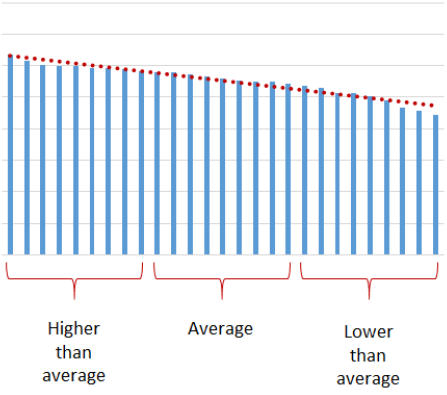
Figure 3: Relationship between infant representation in the admission population and percent of children exiting to permanency within 12 months (sample state)



There may be a perfectly good reason why some counties have proportionally more infants entering care than others, and it may be that in those counties, infants' need for foster care is due to phenomena or policies that are out of the child welfare system's control. As such, when we compare counties' rates of permanency to one another, we don't want to hold age mix "against them." Instead, we want to say: I know that variation in age mix accounts—at least in part—for county-to-county differences in the rate of permanency (diagram A below). If I take away the effect of child age, will the county-level variation in permanency rate go away? If it goes away to the point where there are no longer any significant differences between counties (B), it means that permanency depends much more on child age than on the county where the child lives. If significant differences between counties remain (C), it means that *notwithstanding* the age mix of children in their care, some counties still have a better rate of permanency than others.



Typically, some variation in outcomes remains after controlling for a set of core variables; after adjustment, some counties will be comparatively above average, while others will be below average. Continuing with the example of permanency within 12 months, this means that even after accounting for a set of variables known to influence permanency, some counties will have above average permanency rate (i.e., better performance than expected/over-performing outliers), while others will have a below average permanency rate (i.e., worse performance than expected/under-performing outliers).



Why is knowledge of county outliers important in a CQI context?

The CQI process is a Plan-Do-Study-Act problem solving cycle that demands the use of evidence (see Appendix B). The "Plan" phase of the cycle requires an agency to develop a hypothesis that contains four core statements:

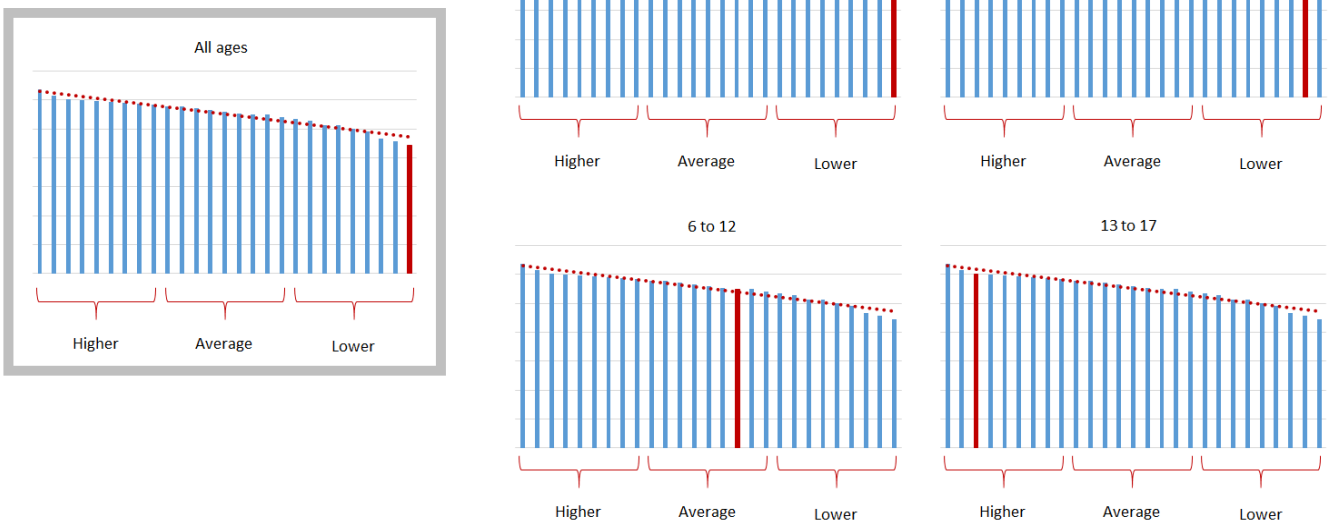
- I observe [some outcome that needs improvement].
- I think it's because [of this reason].
- So I plan to [implement some new way of doing business],
- which I think will result in [an improvement in the outcome of interest].

We can only define an outcome as “in need of improvement” when we can compare it to some level of performance that can be defined as “better.” In other words, **observing variation** is what kick-starts the CQI process. For example: “I observe that the rate of permanency within 12 months in County Z is much lower than it is in other counties and much lower than it is statewide.” That information ignites the notion that County Z could be a worthwhile area for targeting permanency improvement resources.

Risk-adjustment allows us to garner additional information—and potentially, additional confidence—about that strategy. If I can say that **not only** is County Z’s **observed performance** worse than average, but it is **also** significantly worse than average **after** I account for a number of characteristics unique to that county and the children in its foster care system, then I have much more support for the idea that County Z’s permanency rate does in fact need attention.

Breaking down—or **stratifying**—those outcomes by child-level characteristics gives us even further precision. Say we observe that after risk-adjustment, County Z has a permanency rate that is significantly below the state average, highlighted below by the red bar in the box on the left labeled “All ages.” One question is whether that is true for **all children** in County Z or whether the county’s comparatively poor performance is driven by its inability to achieve permanency for a particular age group of children.

Risk-adjusting the permanency outcome **and** stratifying the results by child age gives us the answer. The four charts to the right would suggest that County Z’s below average permanency rate is driven by its below average outcomes for infants and 1 to 5 year olds. When it comes to 6 to 12 year olds, the county performs in the average range after adjustment; when it comes to teens, the county is actually an *over-performer.*



In light of limited resources and the fact that one cannot always expect the same intervention to have the same effect across different sub-populations, it is important for states to target resources to the places where the possibility for improvement is the greatest. Toward that end, the information Hot Spot provides is critical. Hot Spot reports will not tell you decisively where to make investments or what to do in order to improve outcomes—those decisions need to be made in conjunction with other evidence including cost-benefit projections, particularly when under-performing populations are small counties or subgroups. Hot Spot will, however, point you to the populations that do not perform as we would expect them to, all other things being equal.

With that knowledge in hand, administrators can move toward a series of more precise questions aimed at filling in the "I think it's because…" and "So I plan to…" components of the hypothesis development template outlined above. Namely: What policies and practices are shaping performance in the counties that perform worse than we would expect? Is there an intervention that can be targeted to those areas that would improve the outcome? What forces are at play in counties that perform better than we would expect? Are there positive practices or protective factors in those areas that facilitate those better outcomes, and if there are, can we apply lessons from those areas as we develop interventions for poorer performing areas? Is intervention needed for all children or only for a subset of the population?

How can Hot Spot help me respond to the federal CFSR?

In Round 3 of the federal Child and Family Services Review (CFSR), the Administration for Children and Families will also use a risk-adjustment method to compare states to one another on a set of child welfare performance indicators. Hot Spot uses a similar method to compare counties to one another, but there are some important differences. Refer to Appendix A for methodological details.

Perhaps the most important difference between the CFSR outcomes and the information provided by Hot Spot is what the state is supposed to do with the information. The CFSR's risk-adjusted metrics are used to compare states to one another and determine, for each state, the outcomes on which they are performing relatively well, and the outcomes that the state needs to improve. Specifically, for each outcome, states will be classified as either having above average performance, average performance, or below average performance; areas in which the state performs below average will be included in the state's federal Performance Improvement Plan (PIP). In short, the state's results on the CFSR metrics serve as a starting point, identifying for the state the outcomes on it will be expected to improve.

What statewide CFSR outcomes do not offer is guidance regarding where the state should focus—or with whom—as it sets out to improve those outcomes. Think of the PIP as sparking a Plan-Do-Study-Act cycle and think of your state's results on the CFSR metrics as the observation that kicks off the process, for example:

- "I observe that after risk-adjustment, my state performs below the national average on the percent of children exiting to permanency within 12 months."

The next steps of the process require the state to fill in the statements "I think it's because…" and "So I plan to…" Determining where to focus what improvement efforts and with whom requires knowledge about which subpopulations in the state need the most attention. As described in detail above, Hot Spot results can point the state in that direction.

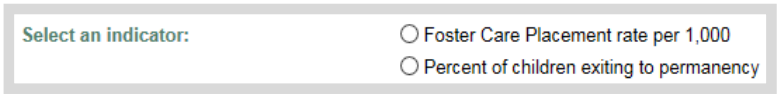
Hot Spot is not a tool for replicating the results of the CFSR formulas (see Appendix A for more information on how to use the FCDA web tool to calculate observed performance and the practical baselines off of which to measure change); rather, it is a tool that helps states make decisions about where to invest resources for improving outcomes for children and families. With regard to the CFSR, this means that Hot Spot assumes the state that has received its PIP requirements and, in response, provides guidance to help the state answer the question: "Now what?"

SELECTING A REPORT

1. Log into the FCDA web tool at https://fcda.chapinhall.org through the **Multistate** portal.

2. On the left hand menu bar, select **Hot Spot**.

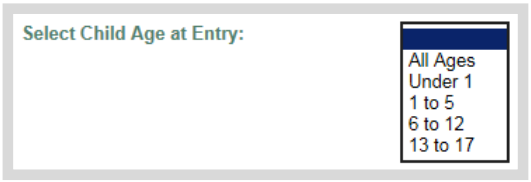
3. On the next page, first select the indicator for which you want to calculate Hot Spots.



- Selecting **Percent of children exiting to permanency** will open a window asking you to select a specific window of time. Choose one.

4. Next, select whether you want to compare county performance to a state or national reference point.

- Selecting **Percent of children exiting to permanency** will open a window asking you to select an age group. Choose one.



5. Click **Submit**.

INTERPRETING THE OUTPUT

The sample below explains output for the **Percent of children exiting to permanency** indicator.

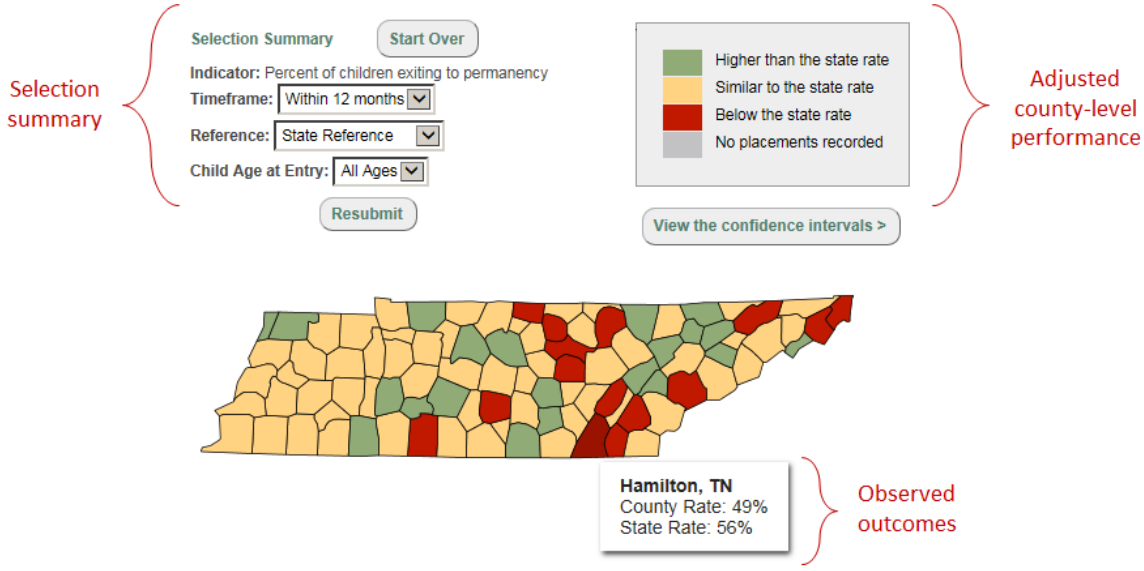
Map output

Selection summary. Your selections will appear at the top of the output page. Adjust these and click **Resubmit** to modify your query.

Observed outcomes. The rollover box shows the observed (unadjusted) county and state/national permanency rates, depending on your selection.

Adjusted county-level performance. Counties are shaded based on the extent to which their permanency rates differ from the reference rate after risk-adjustment.[1] Counties shaded in yellow have permanency rates that are similar to the reference rate. Counties shaded in green are those where the permanency rate is significantly higher than the reference rate; counties shaded in red are those where the permanency rate is significantly lower than the reference rate.

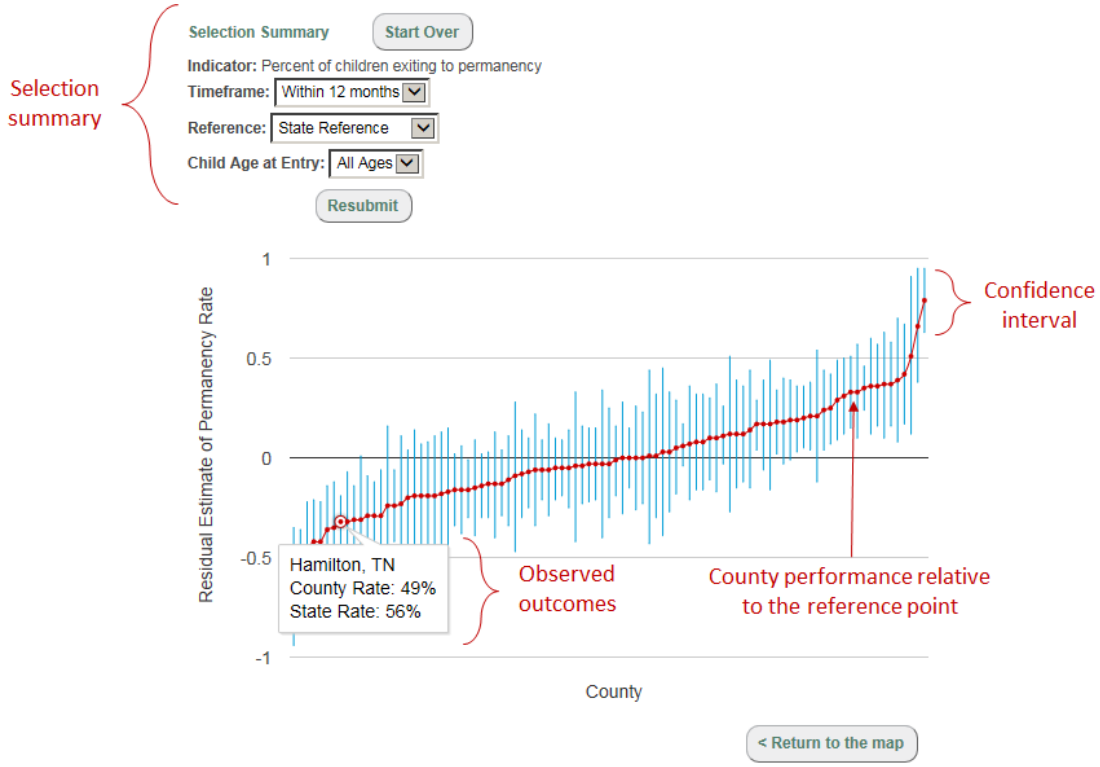
Example. The example below reads as follows: In Hamilton County, the 49% of children entering care exit to permanency within 12 months. Statewide, 56% of children entering care exit to permanency within that timeframe. After adjusting for child- and county-level factors known to influence permanency, Hamilton County's rate of permanency within 12 months is significantly below the statewide rate.



[1] See Appendix A for a list of control variables used in the statistical model.

Confidence intervals

Clicking **View the confidence intervals** on the map output takes you to a page that looks like this:



Selection summary. Again, your selection criteria will appear at the top of the page. Adjust these and click **Resubmit** to modify your query.

Observed outcomes. Each dot represents a different county. Again, the rollover box shows the observed (unadjusted) county and state/national permanency rates.

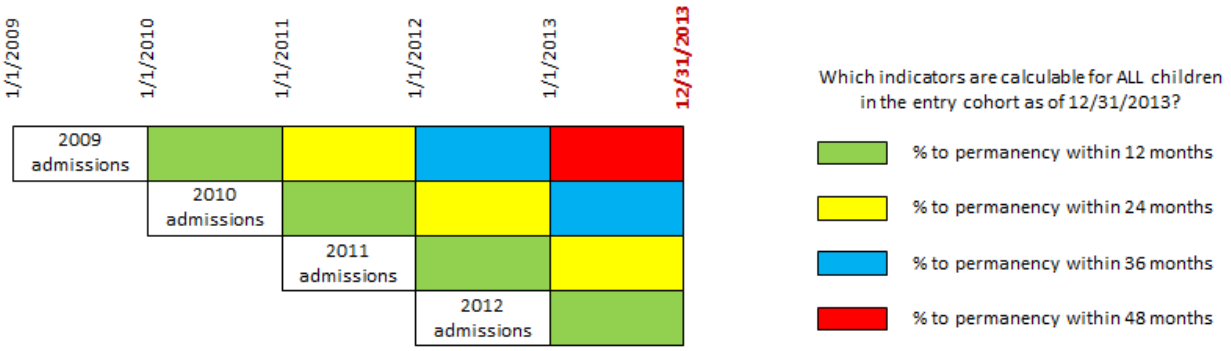
County performance relative to the reference point. The horizontal "zero" line on the vertical axis represents the state/national permanency rate, depending in your selection. The location of the dot represents the extent to which the county permanency rate differs from the reference rate after risk-adjustment.

Confidence interval. The blue bar that passes through the dot is called the confidence interval. Its position indicates whether, after risk-adjustment, the difference between the county rate and the reference rate is statistically significant. If the bar crosses zero, the county rate is **not** significantly different from the reference rate. If the bar falls entirely above the zero line, the county rate is significantly higher than the reference rate; if the bar falls entirely below the zero line, the county rate is significantly lower than the reference rate. The location of the bar with respect to the zero-line corresponds to the shading scheme on the map.

The length of the bar tells you how confident you can be about the difference between the county's expected permanency rate (which is based on the selected reference point—either the state average or the national average) and its observed rate. The length of the bar is affected by a number of factors including the number of children placed in foster care; as a general rule, longer bars will be associated with small counties and short bars will be associated with larger counties. The assessment of whether a county is a statistical over- or under-performer takes the size of the foster care population into account.

A note about censored data

When tracking the experiences of children in an entry cohort, one must bear in mind that it is only possible to observe children's trajectories up until the date as of which the data source is current. Hot Spot, for example, allows the user to calculate the observed and adjusted rates of permanency within 12, 24, 36, and 48 months. The database as of this writing is current—or censored—as of 12/31/2013. As shown in the graphic below, the likelihood of permanency within, say 48 months, is only available for one of the most recent four entry cohorts. We cannot, for instance, fully observe the likelihood of permanency within 48 months for children who entered care in 2010, 2011, or 2012, because as of the censor date, 48 months have not elapsed for all children in those cohorts.



Hot Spot's permanency measure uses a multi-year cohort of admissions between 2007 and 2012. This means that as of 12/31/2013, findings on permanency within 12 months are stable whereas findings on permanency within 24, 36, and 48 months may change as the database is updated. For more on the rationale behind the multi-year cohort, see Appendix A.

APPENDIX A

Hot Spot Methodology

Data source

Hot Spot reports are generated using data from the Multistate Foster Care Data Archive (FCDA)—a longitudinal repository of data on children in foster care in the United States maintained by The Center for State Child Welfare Data. The FCDA is the oldest database of its kind, containing decades of information on nearly 3 million children who have spent time in foster care in more than two dozen states.

The structure of the FCDA is based on foster care spells. A spell is a continuous period of time that a child spends in foster care. A placement event starts a spell; an exit event ends a spell. A child may experience multiple placement events (i.e., movements) during his/her spell. A child may also have more than one spell; this is the case when a child re-enters care after being discharged.

The FCDA contains data from multiple states. In order to maintain a multistate archive that permits valid comparisons between states, the FCDA establishes one set of common inclusion/exclusion criteria:

- Spells are physical custody spells (as opposed to legal custody spells). This means that trial home visits and runaways are classified as exits from foster care. Trial home visits are included in the count of reunifications. Children who are in the legal custody of the state but who do not enter foster care are not included.
- Spells that are fewer than five days long are not included.
- A spell is not included if, at the start of the spell, the child is 18 years old or older.
- If a child exits care by running away, reaching the age of majority, or exiting to a destination classified as “other/unknown” and the child comes back into foster care within seven days of exit, the re-entry is bridged to form one spell, rather than identified as a new spell.
- A spell is forced to exit to “reach majority” on the child’s 21st birthday if the child is still in care at that time.
- If a child exits care to “runaway” and the child’s age at exit is less than 10 years old, the exit type is changed to “other.”
- If a child exits care to “reach majority” and the child’s age at exit is less than 13 years old, the exit type is changed to “other.”

Method for calculating observed and risk-adjusted outcomes

The table below provides detail on how the Hot Spot program calculates each observed outcome and the variables controlled for in the risk-adjustment of each. County and state calculations use county and state data, respectively. “National” calculations use data from all the counties in all of the states that participate in the Multistate FCDA.

	Foster care placement rate per 1,000 children	Percent of children exiting to permanency within 12/24/36/48 months
Denominator	Child population under the age of 18 (2010 census)	All admissions to foster care between CY 2007 and CY 2012.
Numerator	Number of children who entered foster care for the first time between 2007 and 2012 (x 1,000)	Number of children who exited to permanency within 12, 24, 36, and 48 months of entry
Model	Multilevel Poisson count model	Multilevel discrete time hazard model
Data structure	Multistate FCDA spell file with counties as the unit of analysis	Person-period (month) file developed using the Multistate FCDA
Censor date	12/31/2013	12/31/2013
Control variables	County-level: • Socioeconomic index	Child-level variables • Child gender • Child age • Placement type • Spell type County-level: • Socioeconomic index • Random effect estimate of placement rate Person-period indicators

Control variables defined

Control variables are the factors for which the outcomes are adjusted. These variables are sometimes called covariates. Reports describing the results of risk-adjustment analyses, often contain the phrase, "after controlling for _____, we found that…" or "after adjusting for _____, we found that…" or "after holding the effects of _____ constant, we found that…" The variables in the blanks here are the control variables.

- Child gender: This variable accounts for county-level case mix differences in terms of the proportion of males and females entering foster care.

- Child age: This variable accounts for county-level case mix differences in terms of the age of children entering foster care.

- Placement type: This variable accounts for county-level case differences in terms of the types of placements in which children spend 50% or more of their foster care spell (congregate care, kinship care, non-kinship foster care, or "other").

- Spell type: This variable accounts for county-level case mix differences in terms of the proportion of children in the admissions population entering foster care for the first time (as opposed to re-entering care after a previous discharge).

- Socioeconomic index: This variable accounts for county-level differences in socioeconomic disadvantage. The variable is an index calculated using four variables gathered from the United

States 2010 Census: percent of adults that have less than high school education, unemployment rate, percent of female-headed households, and poverty rate.

- Random effect of placement rate captures the fact that placement rates in some counties are higher than in others. Research has shown that permanency rates are affected by placement rates. This adjustment accounts for the county placement rate when comparing permanency rates.

- Person-period indicators are variables that indicate where, temporally, a child is in the course of his/her spell in foster care. They account for the fact that the likelihood of exiting to permanency differs for children who are at different points in their foster care spell.

Why use a multi-year sample?

The risk-adjustment method, like other statistical methods, requires a sufficient sample size in order to produce stable results. This is particularly necessary in the case of small counties and subpopulations where small changes in the number of children achieving a particular outcome can have a dramatic effect on aggregate performance. Therefore, to ensure a sufficient sample size and maximize our confidence in the Hot Spot results, we have chosen to use a large, multi-year sample.

One side effect of this decision is that, because multiple cohorts of children are merged into one, the analysis precludes an assessment of changes in the observed and adjusted outcomes over time (i.e., from year to year).

To see observed outcomes over time, we refer the user to the **Foster Care Profile Report** function of the FCDA web tool, or the **All Spells** page, which enables the user to customize entry cohorts and compare them to one another. These tools, in addition to the **Benchmarks** function, can provide the user with observed outcomes that may be used to develop baselines off of which to measure future change.

Recall that the purpose of Hot Spot is to compare counties to one another and to state and national benchmarks; it is not a tool for examining performance over time. See the **Background** section above for details.

How is the Hot Spot methodology similar to/different from that on the federal CFSR?

Database. As noted above, Hot Spot reports use data from the Multistate FCDA. The FCDA is developed using raw data tables from member states' electronic databases (SACWIS or otherwise). The CFSR uses data from states' AFCARS (and in some cases NCANDS) submissions. These databases are similar but not identical. See above for the inclusion/exclusion criteria used to develop the FCDA.

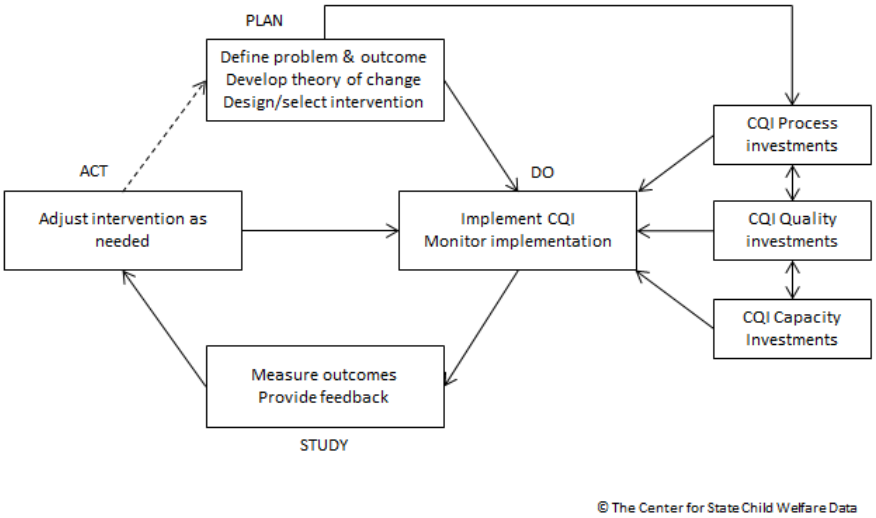
Indicators. Currently, Hot Spot reports out on two main indicators. The first is placement rate per 1,000 children; this is not a CFSR indicator. The second is (2) permanency rate within 12, 24, 36, and 48 months. The CFSR does measure states on their rate of permanency within 12 months; it does not measure state performance with respect to the longer timeframes.

Control variables. In comparing states' rates of permanency within 12 months, the CFSR adjusts for the child's age at entry and the state's foster care entry rate. In examining county performance, Hot Spot adjusts for these factors *in addition* to several others (see table above). We believe that controlling for these additional influences provides a more precise view of over- and under-performing jurisdictions.

Sampling. As noted above, Hot Spot uses a multi-year cohort in order to ensure sufficient sample size. The CFSR uses either a year-long entry cohort, a year-long exit cohort (re-entry), or a point in time sample, depending on the indicator.

APPENDIX B
Evidence use throughout the CQI process[2]

Because CQI models have been used to structure the problem solving process in numerous fields for nearly a century, many variations exist; however, they all contain the same four fundamental phases: Plan, Do, Study, and Act (PDSA). In a child welfare context, these stages unfold as follows:



- **Plan.** The CQI cycle begins when the agency defines a problem it wishes to solve by observing baseline performance on an outcome of interest. Next, the agency identifies an intervention that is expected to improve that outcome and sets targets for improvement. For these purposes, an intervention is any change to the current way of doing business that is expected to bring about a change in outcomes.

- **Do.** Implementing a new intervention requires the agency to invest in one or more of three major areas: the quality of services to be delivered, the processes by which they are delivered, and the capacity of the agency to deliver them with fidelity. Quality and process refer to the "what" and "how" of intervention. Capacity investments are the resources that the agency will allocate to ensure that the intervention is implemented according to process and quality standards. Over the course of the implementation period, the agency conducts process evaluation to monitor the extent to which the intervention is being implemented with fidelity to its design.

- **Study.** After an established period of time, the agency measures the outcome of interest again to determine whether the intervention has had its intended effect.

- **Act.** Finally, the agency uses findings from the process and outcome evaluations to make decisions about its future investments—namely, whether to continue with the intervention, modify or discontinue it, or revisit the original conceptualization of the problem. From there the cycle begins again.

Each stage of the PDSA cycle contains its own hypothesis development/testing activities that must be executed—one must make a claim, make some decision, or take some action. Ensuring that those claims, decisions, and actions are defensible means supporting them with evidence. Each PDSA activity and its associated demands for evidence are outlined in the grid below.

[2] Adapted from Wulczyn, F., Alpert, L., Orlebeke, B., & Haight, J. (2014). *Principles, Language, and Shared Meaning: Toward a Common Understanding of CQI in Child Welfare.* Chicago: The Center for State Child Welfare Data, Chapin Hall at the University of Chicago.

CQI Phase	Hypothesis development/testing	Evidence use
Plan 	Define the problem. ("I observe that…")	What evidence supports this observation?
	Hypothesize as to the cause of the problem. ("I think it's because…")	What evidence supports this theory of change?
	Identify a solution. ("So I plan to…")	What evidence supports the hypothesis that this intervention will have the intended effect on the target population?
	Set a performance target. ("….which I think will result in…")	What evidence supports the hypothesis that the proposed dose of the intervention will lead to this specific degree of improvement?
Do 	Implement the intervention.	Collect data required for an analysis of intervention effectiveness and analysis of implementation fidelity.
	Monitor implementation.	What evidence is there that the intervention was (or was not) implemented with fidelity?
Study 	Measure progress toward the target outcome.	What evidence is there that the intervention was effective (or not effective)?
	Provide feedback to relevant stakeholders and decision makers.	Transmit evidence regarding outcomes and fidelity to those who will interpret the findings and make decisions accordingly.
Act 	Determine the extent to which the problem still exists.	What evidence supports this observation?
	Confirm or refute the theory of change.	What evidence supports this claim?
	Adjust the intervention as needed.	What evidence supports the decision to continue, modify, or discontinue the intervention?