Understanding and measuring child welfare outcomes

Amy C D'Andrade, San Jose State University
K Lemon, San Jose State University
M Austin, San Jose State University
SUMMARY. The new “Children’s and Family Services Reviews” (CFSR) process focuses on the effectiveness of services to children and families by measuring client outcomes. This article reviews the research literature related to child welfare outcomes in order to provide a context for federal accountability efforts. It also summarizes the 2001 federal mandate to hold states accountable for child welfare outcomes and describes California’s response to this mandate. Implications of the outcomes literature review and measurement problems in the CFSR process suggest CSFR measures do not always capture meaningful outcomes. Recommendations for change are made. doi:10.1300/J394v05n01_06

Available online at http://jebsw.haworthpress.com
© 2008 by The Haworth Press. All rights reserved.
doi:10.1300/J394v05n01_06
INTRODUCTION

Efforts to use data to monitor and improve social services are not new. As far back as 1930s, there were calls for accountability for social services (Courtney, Needell & Wulczyn, 2004). More recently, the Government Performance and Reporting Act of 1993 required federal agencies to establish performance goals and monitor performance results for all federal programs (Kautz, Netting, Huber, Borders & Davis, 1997). In addition, the Social Security Amendments of 1994 required the Department to “promulgate regulations for reviews of states’ child and family services” (Administration for Children and Families, n.d.[b]). Finally, the Adoption and Safe Families Act of 1997 required the federal government to develop a set of outcome measures for public child welfare programs (USGAO, 2004).

This article reviews the research literature related to child welfare outcomes in order to provide a context for federal accountability efforts. It also summarizes the 2001 federal mandate to hold states accountable for child welfare outcomes and describes California’s response to this mandate. The federal outcomes and this structured review of the literature focus on client outcomes: namely, outcomes for children as they move in and out of state child welfare systems. The data on the efficiency and effectiveness of specific child welfare programs (e.g., independent living, therapeutic foster care, kinship care, domestic violence or substance abuse treatment) are not included in this review.

The most frequently cited child welfare outcomes in the research literature and in federal and state accountability efforts fall into three broad domains: (1) safety, (2) permanency, and (3) well-being. The outcomes for safety include protecting children from abuse and neglect and maintaining them safely in their own homes. In the permanency domain, outcomes assess whether children in out-of-home care have permanency and stability in their living situations. The outcomes related to well-being include education, physical health, and mental health of children while they are in care and upon emancipation from the system.
OUTCOMES AS REFLECTED IN RESEARCH LITERATURE

The search of the literature for findings related to child welfare safety, permanency and well-being outcomes involve the use of specific search terms for accessing social science and academic databases available through the University of California library. In addition, the search included websites specializing in systematic reviews, publications of research institutes, databases for conference proceedings, dissertation databases, and general internet searches. When child, family or case characteristics have been found to be associated with the outcomes, these are described as well.

Safety Indicators

Child safety is a priority for the child welfare system. The measures of child safety that are assessed in the research literature include: (1) maltreatment recurrence, or the rate at which children experience maltreatment subsequent to an initial investigated event of maltreatment; (2) maltreatment in out-of-home care, or the rate at which children experience maltreatment while placed in foster care; and (3) re-entry to foster care, the rate at which children experience placement into foster care subsequent to reunification with their parents. The research findings related to these indicators are described below.

Maltreatment Recurrence: The findings related to maltreatment recurrence vary depending on the definition of “recurrence” and the time span of the observation period following the initial referral. When “recurrence” is defined as a subsequent referral or report to the child welfare system, studies have found that about one quarter of children experience maltreatment recurrence within 18 months of the initial referral (English, Marshall, Brummel & Orme, 1999; Fuller & Wells, 2003). When “recurrence” is defined as a subsequent substantiated referral or report to the child welfare system, a smaller proportion of referred children experience recurrence within 18 months of the initial referral (English, Marshall, Brummel & Orme, 1999; Lipien & Forthofer, 2004; Terling, 1999) and that proportion grows as more time elapses from the initial referral (Depanfilis & Zuravin, 2002; Depanfilis & Zuravin, 1999; English, Marshall, Brummell & Orme, 1999; Lipien & Forthofer, 2004; Terling, 1999).

The child factors found to be associated with an increased likelihood of child maltreatment recurrence include younger age (Drake, Johnson-Reid, Way & Chung, 2003; Fuller, Wells, & Cotton, 2001; Lipien & Forthofer, 2004; Marshall & English, 1999), health, mental health,
and/or developmental problems (Depanfilis & Zuravin, 2002; Marshall & English, 1999. Additionally, Asian/Pacific Islander children appear to have lower recurrence rates than children of other racial/ethnic backgrounds (Fluke, Yuan & Edwards, 1999). The risk factors related to parents include substance abuse (Fuller & Wells, 2003), criminal history (Fuller & Wells, 2003), domestic violence (Depanfilis & Zuravin, 2002), childhood abuse (Marshall & English, 1999), lack of social support (Depanfilis & Zuravin, 2002), and poverty (Jones, 1998). Families with multiple children (Depanfilis & Zuravin, 2002; Marshall & English, 1999) and single parent-families (Fuller, Wells, & Cotton, 2001) have been found to be more likely than other types of families to have a subsequent substantiated report of child maltreatment. Finally, the risk of maltreatment recurrence increases if the initial report is substantiated (Lipien & Forthofer, 2004) and increases with each subsequent maltreatment incident (Fluke et al., 1999; Fuller et al., 2001; Terling, 1999).

**Maltreatment in Out-of-Home Care:** The federal statistics do not describe what proportion of children in foster or group care nationwide have been maltreated. However, a federal report indicates that less than 1% of perpetrators of maltreatment in 2003 were foster parents or residential staff, with neglect being the most common form of maltreatment reported (U.S. Department of Health and Human Services, 2005). When former foster youth are queried, over 30% report that they experienced some form of child maltreatment while in care; neglect again is the most commonly reported type of maltreatment (Annie E. Casey Foundation, 2005). The studies examining incidence of maltreatment in care have found that between 8 and 120 children per 1000 in care are victims of substantiated maltreatment and the rate varies by placement type (Spencer & Knudsen, 1992).

**Re-Entry to Foster Care:** A substantial portion of children who are reunified with their parents subsequently re-enter care within one to two years. About 13-14% of them re-enter care within one year (Jones, 1998; Needell, Webster, Cuccarro-Alamin, Armijo, Lee, Levy, Shaw, Dawson, Piccus, Magruder, Kim, Conley, Henry, Korinek, Paredes & Smith, 2005), about 20% of them re-enter care within two to three years (Courtney, 1995; Courtney et al., 1997; Festinger, 1996), and the proportion increases as more time elapses since reunification (Frame et al., 2000; Wulczyn, 1991).

A number of child, parent, and case characteristics have been found to be associated with re-entry into care. Infants (Courtney, 1995; Courtney, Piliavin & Wright, 1997; Frame, Berrick, & Brodowski, 2000; Wells & Guo, 1999), African American children (Courtney,
Children with health problems (Courtney, 1995; Jones, 1998) have been found to have a greater likelihood of re-entry. Parents who are poor (Courtney, 1995), who have a history of criminal activity (Frame et al., 2000), substance abuse problems (Frame et al., 2000), or limited social supports (Festinger, 1996) are more likely to have their children re-enter care. Lastly, children placed with kin prior to reunification (Courtney, 1995; Courtney et al., 1997; Frame et al., 2000; Wells & Guo, 1999) and who experience more placement moves while in care (Courtney, 1995; Courtney et al., 1997; Wells & Guo, 1999) are more likely to re-enter care.

**PERMANENCY INDICATORS**

The second primary goal of the child welfare system is permanency; namely, reunifying children with their parents or finding them adoptive homes as quickly as possible. While children remain in care, an important aspect of permanency is the degree of stability they experience in the form of fewer placement changes. The research findings related to permanency indicators are described below.

**Reunification:** Although national data suggest that over half of children exiting care in 2001 were reunified (U.S. Department of Health and Human Services, 2003), rates of reunification in longitudinal studies generally reflect lower rates of reunification that vary between 23-48% of children entering care after 1-2 years (Berrick, Needell, Barth & Jonson-Reid, 1998; Courtney, McMurty & Zinn, 2004; Needell et al., 2005; Wells & Guo, 2003; Wells & Guo, 2004). The rate of reunification varies based on the time period under investigation, with a higher proportion of cases reunifying as more time elapses from entry into care (Barth, 1997; Courtney, 1994; Harris & Courtney, 2003; McMurty & Lie, 1992; Wells & Guo, 1999).

The research on factors affecting reunification have identified a variety of child, family and case characteristics that appear to affect the likelihood of reunification. In general, the research suggests that younger children (Courtney & Wong, 1996; Smith, 2003a), children of color (Courtney & Wong, 1996; Wells & Guo, 1999), and children with health and emotional/behavioral problems (Courtney, 1994; Landsverk, Davis, Ganger, Newton, & Johnson, 1996) are less likely to reunify than children without those characteristics. While poor families are less likely to reunify than those who are not poor (Courtney & Wong, 1996; Smith, 2003a), mov-
ing from welfare to employment also appears to decrease the likelihood of reunification (Wells & Guo, 2003). The children from two-parent homes appear more likely to be reunified than children from one-parent homes (Harris & Courtney, 2003; Wells & Guo, 1999). In terms of parental characteristics, the presence of maternal mental health problems (Wells & Guo, 2004) and homelessness (Courtney, McMurty & Zinn, 2004) decrease the likelihood of reunification. Children initially placed as a result of neglect have been found to be less likely to reunify than children placed for other reasons (Courtney & Wong, 1996; Harris & Courtney, 2003; Wells & Guo, 1999; Wells & Guo, 2003), and children placed with kin reunify more slowly than children placed with non-kin (Courtney & Wong, 1996; Harris & Courtney, 2003). Some studies have found families receiving services are more likely to reunify than those not receiving these services (Courtney & Wong, 1996; Smith, 2003a).

Adoption: The national data indicate that among children exiting care in 2001, 18% were adopted (U.S. Department of Health and Human Services, 2003). The national data also suggest that a sizable portion of children wait long periods in out-of-home care before adoption. The research using longitudinal data has generally found lower adoption rates than those reported in federal exit cohort data, with rates varying from about 2% after 2-3 1/2 years (Courtney, 1994 Needell et al., 2005; Berrick et al., 1998), to 9-20% after 6 years (Barth, 1997; Berrick et al., 1998; McMurty & Lie, 1992). The child characteristics associated with a decreased likelihood of adoption include male gender (Kemp & Bodonyi, 2002; Kemp & Bodonyi, 2000), younger age (Barth, 1997; Courtney & Wong, 1996; Kemp & Bodonyi, 2002; Smith, 2003b), non-white ethnicity (Barth, 1997; Courtney & Wong, 1996; Kemp & Bodonyi, 2002; Smith, 2003b), health problems or disabilities (Courtney & Wong, 1996; Smith, 2003b), and placement with kin (Courtney & Wong, 1996; Smith, 2003b). Some research suggests demographics such as urban residence (Courtney & Wong, 1996) and state of residence (Smith, 2003b) may affect the likelihood of adoption as well.

Placement Stability: The studies of placement stability often use different definitions of stability and different time periods for observation. In general, research suggests that the more time children spend in out-of-home care, the more placements they experience. After 1-2 1/2 years, about 20-40% of children still in care experience three or more placements (Berrick et al., 1998; Palmer, 1996; Pardeck, 1984; Needell et al., 2005) and after 3-4 years, about 40-50% of children still in care
have had three or more placements (Berrick et al., 1998; Fernandez, 1999; Usher, Randolph, & Gogan, 1999). There are a number of factors associated with placement disruptions, including such child factors as male gender and African American ethnicity (Webster, Barth & Needell, 2000), older age (James, Landsverk & Slyman, 2004; Smith, Stormshak, Chamberlain & Whaley, 2001; Webster et al., 2000; Wulczyn, Kogan & Harden, 2003), and child behavior problems (Barber, Delfabbro & Cooper, 2001; Newton, Litrownik, & Landsverk, 2000; Palmer, 1996). Children placed as a result of neglect (Barber et al., 2001; Webster et al., 2000; Wulczyn et al., 2003) and children placed with kin (Webster et al., 2000; Wulczyn et al., 2003) tend to have more placement stability than children without those characteristics.

**Well-Being Indicators**

Enhancing child and family well-being is a third goal of the child welfare system. Physical health, mental health and educational problems among children in the child welfare system have been fairly well documented in the research literature, although differing research methodologies present some challenges in interpreting findings. Many studies use data collected at one point in time, which may over-sample children who have been in the child welfare system for long periods, possibly inflating rates of mental, physical and educational problems. Some studies suggest that children come into the child welfare system with numerous problems and it is often difficult to determine if problems are improved or exacerbated by experiences in out-of-home care.

*Physical and Mental Health Issues:* Children entering the child welfare system appear to have a number of physical health problems (Chernoff, Combs-Orme, Risley-Curtis & Heisler, 1994; Hochstadt, Jaudes, Zimo, & Schachter, 1987), in addition to relatively high rates of developmental delays (Chernoff et al., 1994; Leslie, Gordon, Ganger & Gist, 2002) and emotional and behavioral problems (Clausen, Landsverk, Ganger, Chadwick & Litrownik, 1998; Halfon, Berowitz & Klee, 1992; Harman, Childs, & Kelleher, 2000; Landsverk, Davis, Ganger, Newton, & Johnson, 1996; McIntyre & Kessler, 1986).

*Educational Issues:* The research consistently notes the educational deficits among children in foster care. A substantial portion of these children have repeated a grade, and/or receive SED services (Chernoff et al., 1994; Flynn & Biro, 1998). Children in the child welfare system have been found to be more likely than other children to have low levels of engagement in school, to be suspended or expelled, to change schools, and to receive lower grades (Eckenrode, Laird & Doris, 1993;
Preparation for Independent Living. Annually, approximately 20,000 youth are discharged from the foster care system to “independent living” (U.S. General Accounting Office, 1999). Available research suggests that foster youth who “age out” of the system face serious challenges, such as difficulty accessing health insurance and mental health services (Courtney, Piliavin, Grogan-Kaylor, & Nesmith, 2001; Merdinger, Hines, Lemon, & Wyatt, in press; Reilly, 2003), incarceration (Courtney et al., 2001; Reilly, 2003), housing instability and homelessness (Cook, 1994; Courtney et al., 2001), and low high school completion/GED rates (Barth, 1990; Blome, 1997; Cook, 1994; Courtney et al., 2001; Festinger, 1983; Mech, 1994; Reilly, 2003; Zimmerman, 1982).

THE FEDERAL REVIEW PROCESS

While previous federal review and accountability processes focused almost entirely on the accuracy and completeness of case files and other records, the new federal “Children’s and Family Services Reviews” (CFSR) process focuses on the effectiveness of services to children and families by measuring client outcomes. The CFSR process was launched in 2001; all 50 states, plus the District of Columbia and Puerto Rico, have now completed their CSFR reviews.

The review process has three phases. First, administrative data are summarized to assess certain quantitative indicators for each state. Second, an on-site review is conducted of a sample of 50 cases (half are foster care cases, and half in-home services cases) from three sites (Administration for Children and Families, May, 2002). Reviewers spend one week reviewing cases and interviewing agency stakeholders (such as judges or advocates) and case-specific stakeholders (such as parents, workers, and children) (U.S. General Accounting Office, April 2004) in order to determine whether each case is in “substantial conformity” with seven overall outcomes (U.S. Department of Health and Human Services, 2003). If the state is found to be out of compliance on any of the outcomes based on both the administrative data and the on-site review process, the third phase involves the development of a program improvement plan. After a two-year implementation period, changes in the outcomes are assessed. If agreed upon targets have not been met by that time, financial penalties are assessed (Administration for Children and Families, August 2001).
A total of 26 different indicators are used to assess the seven outcomes. Of these indicators, 3 rely on the administrative data only, 20 rely on the on-site data only, and 3 rely on both the on-site review and administrative data sources. Figure 1 provides a summary of the federal CSFR outcomes, the indicators used to measure each outcome, and the sources of information for evaluating the indicator.

The federal government has established the minimum performance level that a state must attain in order to be in “substantial conformity” with the outcomes. For outcomes based solely upon administrative data, a state must meet or exceed the standard established by the federal government. The standards are set at the point at which approximately 25% of states had performed better and 75% had performed worse in AFCARS and NCANDS submissions (Administration for Children and Families (a); Courtney, Needell & Wulczyn, 2004). Figure 2 displays the measures for the six administrative data indicators as well as the national standards. For outcomes based solely upon on-site case review data, 90% of cases reviewed in the state must be found to be in “substantial conformity.” For those outcomes based on both on-site reviews and administrative data, both requirements must be met.

No state has achieved substantial conformity on all the outcomes. Figure 3 shows the number and proportion of jurisdictions achieving substantial conformity on the seven outcomes. California did not meet any of the national standards for the administrative data indicators, and was not in substantial conformity with any of the seven outcomes. As of January 2004 no penalties had been applied, but potential penalties range from $91,492 for North Dakota to $18,244,430 for California (U.S. General Accounting Office, April 2004).

**Measurement Issues**

The federal government and many state officials report that the CSFR process is valuable. In the 2004 GAO survey, 26 of 36 responding states either generally or completely agreed with results of their final CSFR report, even though none of the states achieved substantial conformity with all the outcomes. As a result of the process, some states reported improved relationships with community stakeholders, as well as increased public and legislative attention being given to important child welfare issues (USGAO, 2004).

However, a number of measurement issues regarding the federal outcomes have been raised. State officials in all five states visited by the GAO office in 2004 expressed concerns that AFCARS and NCANDS
FIGURE 1. Federal Outcomes, Indicators, and Data Source

<table>
<thead>
<tr>
<th>Domain</th>
<th>Outcome</th>
<th>Indicator</th>
<th>Case Reviews</th>
<th>Admin Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAFETY</td>
<td>Children are protected from abuse and neglect</td>
<td>1. Timeliness of investigations of reports</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Recurrence of maltreatment</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Incidence of abuse or neglect in foster care</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children are safely maintained in their homes</td>
<td>4. Services to family to protect children/prevent removal</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. Current risk of harm to child</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>PERMANENCY</td>
<td>Children have permanency and stability in their living arrangements</td>
<td>6. Foster care re-entries</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7. Stability of foster care placement</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8. Permanency goal for child</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9. ILS (2001); reunification, guardianship or permanent placement with relative (2002-2004)</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>10. Achievement of adoption</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>11. Permanency goal of &quot;other planned living arrangement&quot;</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>12. Time to reunification</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>13. Time to adoption</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Continuity of family relationship is preserved</td>
<td>14. Proximity of current placement</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>15. Placement with siblings</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>16. Visiting with parents and siblings</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>17. Relative placement</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>18. Current relation of child in care with parents</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>19. Preserving connections</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>WELL-BEING</td>
<td>Families have enhanced capacity to provide for children's needs</td>
<td>20. Needs and services of child, parents, foster parents</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>21. Child and family involvement in case planning</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>22. Worker visits with child</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>23. Worker visits with parents</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children receive appropriate services to meet educational needs</td>
<td>24. Educational needs of child</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>25. Physical health of child</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children receive adequate services to meet their physical and mental health needs</td>
<td>26. Mental health of child</td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>

data, upon which administrative data indicators are based, were not reliable. In addition, researchers have argued that administrative and case review data indicators may not be good measures of the phenomena of interest.

**Administrative Data Indicators:** The administrative data indicators have a number of measurement problems. First, these indicators do not capture important aspects of child welfare processes, such as the rate of
FIGURE 2. Administrative Data Indicator Measures and National Standards

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Measurement</th>
<th>National Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recurrence of maltreatment</td>
<td>Of all victims of substantiated child abuse or neglect during the first six months of the period under review, what % had another substantiated or indicated report within 6 months.</td>
<td>6.1%</td>
</tr>
<tr>
<td>Incidence of abuse or neglect in foster care</td>
<td>For all children in foster care during the period under review, what % were the subject of substantiated or indicated maltreatment by a foster parent or facility staff.</td>
<td>0.57%</td>
</tr>
<tr>
<td>Foster care re-entries</td>
<td>For all children who entered foster care during the year under review, what % of them re-entered care within 12 months of a prior episode.</td>
<td>8.6%</td>
</tr>
<tr>
<td>Stability of foster care placement</td>
<td>Of all children who have been in foster care less than 12 months from the time of the latest removal, what % had no more than 2 placement settings.</td>
<td>86.7%</td>
</tr>
<tr>
<td>Time to reunification</td>
<td>Of all children reunified with their parents at time of discharge from foster care, what % were reunified in less than 12 months from the time of the latest removal from home.</td>
<td>76.2%</td>
</tr>
<tr>
<td>Time to adoption</td>
<td>Of all children who exited foster care during the year under review to a finalized adoption, what % did so in less than 24 months from the time of the latest removal from home.</td>
<td>32.0%</td>
</tr>
</tbody>
</table>

reunification and adoption. None of the six indicators relate to family and child well-being or to emancipated youth. Similarly, some do not capture the experience of important subsets of children. For example, placement stability is a far greater problem for youth who have been in care for longer periods, yet the related indicator captures the phenomenon only for children in care for 12 months or less.

Second, the indicators do not take into account the dynamic nature of the child welfare system. Changes in one outcome can affect other outcomes (Courtney, Needell & Wulczyn, 2004; Goerge, Wulczyn & Harden, 1996; Tilbury, 2004; Usher, Wildfire & Gibbs, 1999; Wells & Johnson, 2001). For example, decreasing the time to reunification is problematic if the re-entry rate increases as a result. Outcomes need to be considered in the context of other outcomes.

Third, the indicators do not take into account differences between states. According to Goerge et al., "...states exhibit a rather stunning degree of diversity..." (Goerge, Wulczyn & Harden, 1996, p. 25). These differences can include caseload dynamics (caseload population counts), use of kin placements, rate of entry, racial/ethnic populations, poverty, ethnicity, age and other variables that are likely to influence the outcomes. However, all states are required to meet the national standards, regardless of these differences.

Fourth, the indicators are limited by the format of the datasets from which they are drawn, and do not capture longitudinal caseload dynamics. As a result, indicators that require a longitudinal view, such as re-entry, cannot be adequately captured. Currently, the re-entry indicator represents the portion of current entries to care that are re-entries, a
statistic that does not convey information about the rate at which cases re-enter care.

Fifth, several indicators rely upon exit cohorts to describe case phenomena. Exit cohorts are likely to be biased in important ways, since they exclude all youth who do not leave care. As a result, indicators derived from exit cohorts will tend to misrepresent the proportion of cases achieving permanency outcomes within the time frames (Courtney, Needell & Wulczyn, 2004). Exit cohorts are also heavily influenced by population dynamics, such as the number of children entering or exiting care per year. When these dynamics shift, length of stay estimates based on exit cohorts will change as well, even if nothing in the system has occurred that would affect them (Wulczyn, Kogan & Dilts, 2001). These problems are intensified when indicators based upon exit cohorts are used to measure change over time. Research studies have demonstrated that performance trends differ markedly according to whether an entry or an exit cohort is used to assess change, even occasionally heading in opposite directions (Courtney, Needell & Wulczyn, 2004).

Lastly, there are concerns regarding the amount of improvement the federal government will be requiring states to make on the administrative indicators in order to avoid financial penalties. To determine how much states should be required to improve on each administrative indicator, the federal government treated the data submissions of the 52 jurisdictions as a sample, then derived the “sampling error.” This sampling error is the amount by which states must improve. However, the variability within the 52 jurisdiction sample is likely to be substantially greater than the variability of an individual state’s performance over time, particularly if the state is large. Applying the sampling error derived from the 52 sample to every individual state is inappropriate and places a much greater burden upon larger states.

On-Site Review Indicators: A primary concern regarding the on-site reviews is the small sample size of 50 cases, half of which are in-home services cases and half foster care cases. While small samples can sometimes adequately reflect patterns that exist in a population, this is likely only when the sample is randomly selected. Moreover, because not every one of the cases in the sample has relevance for each indicator assessed in the on-site review, sometimes as few as one or two cases are used to evaluate a state’s performance (USGAO, 2004). For example, in Wyoming only 2 cases were relevant to assess the on-site indicator of time-to-adoption. In one of these cases, reviewers determined that appropriate efforts had not been made to achieve the outcome. As a result, the state was assessed as “needing improvement” in this area (USGAO,
In California, 49 cases from three sites—Los Angeles, San Mateo, and Stanislaus—represented over 100,000 children receiving services in California.

A second concern is that in spite of the small sample, data from the on-site record reviews and interviews are heavily weighted in the CSFR process: 23 of the 26 indicators are based upon data from on-site reviews. Additionally, impressions arising from interviews and focus groups may be distorted when some participants are more vocal, even if the experiences they describe are not common. According to a state official in Arizona, one vocal participant in a focus group or interview can have an unreasonably large effect. “Those single comments too often become part of the case (review) report” (Stack, 2005, p.18).

California’s Accountability Efforts

California passed legislation AB 636 in 2001 in response to both the federal outcomes reporting requirements and the limitations of the indicators as performance measures. The “Child Welfare System Improvement and Accountability Act” of 2001 introduces an accountability system designed to facilitate continuous improvements in each county. Beginning in January 2004, “California Child and Family Service Reviews” were initiated in each of California’s 58 counties. These include a set of administrative performance indicators (see Figure 4). While a subset of these parallel the federal CFSR administrative data indicators, another subset goes beyond the federal effort by using California’s own database, the Child Welfare Services Case Management System (CWS/CMS). CWS/CMS data are shared with the Center for Social Services Research at the University of California at Berkeley where it can be reconfigured and analyzed longitudinally. This longitudinal database can be used to generate outcomes that reflect the performance of the system and changes in that performance over time (CDSS, n.d.).

The California and federal accountability efforts differ in several important respects. First, the California approach is more comprehensive, utilizing more administrative data indicators including measures related to well-being and emancipating youth. Secondly, these measures are more carefully constructed. For example, the federal indicator assessing maltreatment recurrence includes all children who experienced an initial referral. However, children who were removed at the time of the initial referral are much less likely to experience a subsequent referral, as they are now in state custody; therefore, one California indicator related to this area excludes these children from consideration. Third, California’s data are con-
FIGURE 3. Number and Proportion of States Achieving Substantial Conformity on Outcomes

<table>
<thead>
<tr>
<th>Domain</th>
<th>Outcome</th>
<th>% of states</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety</td>
<td>Children are protected from abuse and neglect</td>
<td>6</td>
<td>12%</td>
</tr>
<tr>
<td></td>
<td>Children are safely maintained in their homes</td>
<td>6</td>
<td>12%</td>
</tr>
<tr>
<td>Permanency</td>
<td>Children have permanency and stability in their living arrangements</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Continuity of family relationship is preserved</td>
<td>7</td>
<td>14%</td>
</tr>
<tr>
<td>Well-Being</td>
<td>Families have enhanced capacity to provide for children’s needs</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Children receive appropriate services to meet educational needs</td>
<td>16</td>
<td>31%</td>
</tr>
<tr>
<td></td>
<td>Children receive adequate services to meet their physical and mental health needs</td>
<td>1</td>
<td>2%</td>
</tr>
</tbody>
</table>

figured longitudinally, allowing accurate estimates of outcomes like re-entry to foster care. And fourth, indicators assessing the proportion of cases attaining permanency outcomes within certain time frames are based upon entry cohorts. Entry cohorts provide better estimates of change over time than do exit cohorts.

Lastly, the state did not establish any particular standards and counties are not expected to meet a particular performance goal (with the single exception of the measure monthly worker visits with children that requires a level of 90% compliance). Instead, counties identify areas for improvement based on their performance on the measures. To enhance their understanding of problem areas, counties conduct “peer quality reviews.” Relevant cases are randomly selected and interviews are conducted with the social workers involved with the case, clients, and other personnel. This process generates qualitative information that “provides an in-depth analysis of case results and promotes information sharing that helps build the capacity of social workers and other staff” (CDSS, n.d.). This strategy eliminates direct comparisons of outcomes between counties that may have very different economic and demographic characteristics.

**IMPLICATIONS AND RECOMMENDATIONS**

This review of the child welfare research literature provides a context for assessing federal and state measurement and accountability efforts. However, researchers and federal administrators have framed outcomes differently. While federal reports and outcomes use exit cohorts to de-
FIGURE 4. California 636 Administrative Indicators

<table>
<thead>
<tr>
<th>Area</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety</td>
<td>Of all children with substantiated allegation within first 6 months of study period, what % had another substantiated allegation within 6 months? (Federal indicator #2)</td>
</tr>
<tr>
<td></td>
<td>Of all children with a substantiated allegation during the 12 month study period, what % had a subsequent substantiated allegation within 12 months?</td>
</tr>
<tr>
<td></td>
<td>Of all children with a first substantiated allegation during the 12 month study period, what % had a subsequent substantiated allegation within 12 months?</td>
</tr>
<tr>
<td></td>
<td>Of all children with an inconclusive or substantiated allegation during the 12 month period who were not removed, what % had a subsequent substantiated allegation within 12 months?</td>
</tr>
<tr>
<td></td>
<td>Of all children in foster care, what % had substantiated allegation by a foster parent? (Federal indicator #3)</td>
</tr>
<tr>
<td></td>
<td>What % of child abuse and neglect referrals in the study quarter have resulted in an in-person investigation [stratified by immediate and 10 day]?</td>
</tr>
<tr>
<td></td>
<td>Of all children who required a monthly social worker visit, how many received them?</td>
</tr>
<tr>
<td>Permanency</td>
<td>For all children who entered foster care during the year under review, what % of them re-entered care within 12 months of a prior episode? (Federal indicator #6)</td>
</tr>
<tr>
<td></td>
<td>For all children entering foster care for the first time and staying in care for 3 or more days during the 12 month period, and reunified within 12 months of entry, what % re-entered care within 12 months?</td>
</tr>
<tr>
<td></td>
<td>Of all children who have been in foster care less than 12 months from the time of the latest removal, what % had no more than 2 placement settings? (Federal indicator #7)</td>
</tr>
<tr>
<td></td>
<td>For all children entering foster care for the first time and staying in care for 3 or more days during the 12 month period, and were in care for 12 months, what % had no more than 2 placements?</td>
</tr>
<tr>
<td></td>
<td>Of all children reunified with their parents at time of discharge from foster care, what % were reunified in less than 12 months from the time of the latest removal from home? (Federal indicator #12)</td>
</tr>
<tr>
<td></td>
<td>Of all children entering foster care for the first time and staying in care for 5 or more days during the 12 month study period, what % were reunified within 12 months?</td>
</tr>
<tr>
<td></td>
<td>Of all children who exited foster care during the year under review to a finalized adoption, what % did so in less than 24 months from the time of the latest removal from home? (Federal indicator #13)</td>
</tr>
<tr>
<td></td>
<td>Of all children entering foster care for the first time and staying in care for 5 or more days, what % were adopted within 24 months?</td>
</tr>
<tr>
<td>Well-being</td>
<td>For all children in care at the point-in-time of interest, of those with siblings in care, what % were placed with some or all siblings [stratified by all/some]?</td>
</tr>
<tr>
<td></td>
<td>For all children entering foster care for the first time (5 days+) during the 12 month study period, what % were in each placement type? [stratified by first placement, predominant placement, point-in-time]?</td>
</tr>
<tr>
<td></td>
<td>Of those children identified as American Indian, what % were placed with relatives, non-relative Indian, and non-relative Indian families?</td>
</tr>
</tbody>
</table>

termine the proportion and timelines of cases that reunify or are adopted, researchers have not used this sampling strategy due to the biases involved. This makes it difficult to assess whether the national standards are reasonable in the context of the historical achievements of the system. However, a number of conclusions can be drawn from the research literature on child welfare outcomes.

First, there is clearly plenty of room for improvement in child welfare outcomes, and the federal government's effort to assess outcomes is an important step in the right direction. Second, some of the important outcomes that researchers have been studying over the last few decades are
not captured by the current federal administrative data outcome indicators (e.g., the proportion of cases overall that reunify or are adopted, or placement stability for children in long-term care). Third, a myriad of factors appear to influence each outcome, suggesting that comparisons between states could be misleading if these factors are not taken into account. And fourth, while the outcomes of youth in care and emancipating from the system related to well-being are generally poor, this area is not emphasized in the federal review process.

Additionally, the measurement problems in the federal review process have several implications. First, the distortion from using estimates based upon exit cohorts (combined with the questionable reliability of the data from the on-site reviews due to the small sample size) suggest that conclusions about state performance drawn from these data sources could very well be erroneous. As a result, heavy fines could be levied inappropriately. The potential consequences for California are substantial; the state stands to lose more than 18 million dollars, more than any other state (USGAO, 2004).

Secondly, because the understanding gained from these data could be inaccurate, "corrective action" taken by a state to improve outcomes could negatively affect the true outcomes being sought (Courtney, Needell & Wulczyn, 2004). Since financial penalties will be imposed if targets are not met, states have a strong incentive to achieve the targets even if these efforts do not necessarily serve the best interests of children and families (Courtney, Needell & Wulczyn, 2004). For example, in order to reach the re-entry target, an agency might reunify fewer families, since fewer reunified families results in fewer re-entries. Similarly, current practices that benefit children might negatively affect the outcomes (USGAO, 2004). For example, successful efforts to move children currently in long-term foster care into adoptive homes would negatively affect a state’s performance on the adoptions indicator as currently defined; any child adopted after having been in care over 24 months will reduce the proportion of those adoptions that are completed within 24 months.

With the CSFR review process, the federal government has chosen to hold states accountable for what can be counted, even though these measures do not always capture meaningful outcomes. To correct the situation, three changes related to administrative indicators are recommended: First, administrative indicators should be redefined based upon entry cohorts and longitudinal data, rather than exit cohorts and point-in-time samples, so that a more accurate depiction of case processes can be obtained. Second, additional administrative data indica-
tors (based upon longitudinal entry cohorts) should be incorporated into the review process in order to capture important aspects of child welfare case, such as the proportion of cases reunified, adopted, and still in care at certain points in time). Third, national standards for administrative indicators should be eliminated. Given the diversity in states' characteristics, they should only be compared against themselves. If this is not possible, estimates could be risk-adjusted. For example, while incorporating all relevant risk factors would be impossible, it would not be difficult to use some basic demographics like age and race to adjust performance estimates (Courtney, Needell & Wulczyn, 2004).

Additionally, states should ensure their data systems allow for a longitudinal view of child welfare cases. While the changes to SACWIS systems that would be necessary to facilitate this change may involve some costs to states, they would not be difficult to undertake (Courtney, Needell & Wulczyn, 2004). States would also be well-advised to develop their own accountability systems based upon longitudinal data in order to better understand their own performance and make corresponding program and policy adjustments as well as be prepared to defend their performance should findings from the federal CSFR process differ from their own assessments.

The measurement concerns regarding the administrative indicators arise from the limitations of AFCARS and NCANDS data. These databases do not link files for children from year to year, a structure that does not allow a longitudinal consideration of children's experiences (Courtney, Needell & Wulczyn, 2004). Ultimately, AFCARS and NCANDS datasets need to be overhauled so that the federal government can gain more accurate understanding of state processes and achievements (Courtney, Needell & Wulczyn, 2004). Until AFCARS and NCANDS are reconstituted, states should be allowed to utilize other data sources in their CFSR assessments and these should be considered before final CFSR determinations are made (USGAO, 2004).

On-site case review and interview data should not be used to assess state performance, unless a true random sample of a reasonable size can be drawn. If this is not possible, a small, non-random sample might be useful as a way to explore possible explanations for outcomes seen in administrative data.

Lastly, federal and state legislatures need to devote resources to helping public child welfare agencies carry out their responsibilities for accountability (Courtney, Needell & Wulczyn, 2004). States need the ability to configure data so that it conveys meaningful information for management and accountability efforts. This requires the resources to
hire personnel with the capacity to conceptualize and calculate appropriate measures of systems improvements. These resources are needed so that states can evaluate and improve the outcomes of services to children and families.

REFERENCES


doi:10.1300/J394v05n01_06