Effects of Enhanced Foster Care on the Long-term Physical and Mental Health of Foster Care Alumni

Ronald C. Kessler, PhD; Peter J. Pecora, PhD; Jason Williams, MS; Eva Hiripi, MA; Kirk O’Brien, PhD; Diana English, PhD; James White, PhD; Richard Zerbe, PhD; A. Chris Downs, PhD; Robert Plotnick, PhD; Irving Hwang, MA; Nancy A. Sampson, BS

Context: Child maltreatment is a significant risk factor for adult mental disorders and physical illnesses. Although the child welfare system routinely places severely abused and/or neglected children in foster care, no controlled studies exist to determine the effectiveness of this intervention in improving the long-term health of maltreated youth.

Objective: To present results of the first quasi-experimental study, to our knowledge, to evaluate the effects of expanded foster care treatment on the mental and physical health of adult foster care alumni.

Design: We used a quasi-experimental design to compare adult outcomes of alumni of a model private foster care program and 2 public programs. The latter alumni were eligible for but not selected by the private program because of limited openings. Propensity score weights based on intake records were adjusted for preplacement between-sample differences. Personal interviews administered 1 to 13 years after leaving foster care assessed the mental and physical health of alumni.

Setting/Participants: A representative sample of 479 adult foster care alumni who were placed in foster care as adolescents (14-18 years of age) between January 1, 1989, and September 30, 1998, in private (n = 111) or public (n = 368) foster care programs in Oregon and Washington. More than 80% of alumni were traced, and 92.2% of those traced were interviewed.

Intervention: Caseworkers in the model program had higher levels of education and salaries, lower caseloads, and access to a wider range of ancillary services (eg, mental health counseling, tutoring, and summer camps) than caseworkers in the public programs. Youth in the model program were in foster care more than 2 years longer than those in the public programs.

Results: Private program alumni had significantly fewer mental disorders (major depression, anxiety disorders, and substance use disorders), ulcers, and cardiometabolic disorders, but more respiratory disorders, than did public program alumni.

Conclusion: Public sector investment in higher-quality foster care services could substantially improve the long-term mental and physical health of foster care alumni.

Arch Gen Psychiatry. 2008;65(6):625-633

Child abuse and neglect have well-documented long-term effects on adult mental illness.\(^1\)\(^2\) Increasing evidence is emerging that child abuse and neglect also have significant effects on adult physical disorders such as ischemic heart disease,\(^3\) liver disease,\(^4\) and lung cancer,\(^5\) presumably mediated by health risk behaviors.\(^6\)\(^7\) Because of their diverse adverse effects, child abuse and neglect are important targets for preventive interventions. First-line responses typically include family treatment and home-based service programs aimed at reducing maltreatment while keeping the child in the home.\(^8\)\(^9\) However, in cases where the maltreatment is deemed by the child welfare system to be so severe that in-home intervention is considered infeasible, the child is removed from the home and placed in foster care. Therapeutic or group treatment placements are used when behavior problems are a main reason for placement or when the child is thought to be so disruptive or emotionally impaired that a family foster placement is infeasible.\(^10\) In the much more typical case, however, where youth are placed in care primarily because of family maltreatment, the child is placed in family foster care with a relative or an unrelated foster family. Some sense of scope can be obtained by noting that more than 3 million reports of child abuse and neglect are investigated by the child welfare system in the United States each year and, of these, some 900,000 children are classified as maltreated.\(^11\) Approximately 600,000 of these cases are addressed ex-
clusively with in-home treatment programs, whereas the remaining 300,000 children are removed from their homes and placed in foster care. More than 800,000 US children (close to 1% of all US children) are in foster care each year and more than 500,000 are in foster care on any given day of the year.11-13

The child welfare system is financed by a patchwork of government resources totaling more than $24 billion per year14 plus additional private resources.15 Still, reports are common that caseworkers have excessive caseloads and low salaries and that ancillary resources are inadequate, resulting in high staff turnover and uneven job performance.16,17 Calls to redesign the child welfare system to improve care have been made.18 A number of private agencies have developed model foster care programs that offer expanded services, lower caseloads, and higher salaries for workers. However, no systematic evidence exists regarding whether these model programs make a difference to youth outcomes. We herein report data from the first study, to our knowledge, to provide information on the long-term health effects of a model foster care program. We used a quasi-experimental design to compare the mental and physical health of adult foster care alumni interviewed up to 13 years after leaving care from the Casey Family Programs (hereinafter referred to as Casey or the Casey program), a model private foster care program in Oregon and Washington, and from the public foster care systems in these states. To make the comparisons meaningful, we limited the public system alumni to those eligible for but not selected by the Casey program because of limited openings.

METHODS

FOSTER CARE PROGRAMS

The sample included adult foster care alumni of the Casey program; the Oregon Department of Human Services: Children, Adults, and Families Division (hereinafter referred to as Oregon); and the Washington Department of Social and Health Services, Children’s Administration, Division of Children and Family Services (hereinafter referred to as Washington). Casey is a private foster care agency supported by a large endowment established by Jim Casey, cofounder of United Parcel Service, and supplemented by state contracts. Although a detailed description of the unique features of the Casey program compared with standard public foster care programs is beyond the scope of this report, perhaps the most critical feature is cost.

At the time of the intervention evaluated herein, the Casey program had a cost per child roughly 60% higher than that of the public programs ($62 per child per day in 1998 compared with $49 in Oregon and $51 in Washington). This higher cost is due to inclusion in the Casey program of a number of features and services not available in the public programs.19 In brief, Casey caseworkers have higher levels of education than public foster care caseworkers (98% with masters’ degrees vs 36%-42% in Oregon and Washington), lower caseloads (15-17 vs 25-31 cases in Oregon and Washington), higher salaries, and access to a wider range of ancillary services for youth (eg, mental health counseling, tutoring, and summer camps). In addition, Casey foster parents during the study period were paid a $100 per month retainer not available to public program foster parents, were given more financial resources to provide for the foster child, and had access to more case manager assistance. Turnover of foster parents and caseworkers was substantially lower in the Casey program than in the public programs, resulting in more stable placements.

Perhaps one of the most noteworthy features of the Casey program is that it is the largest model foster care program in the country to offer postsecondary job training or college scholarships (tuition, room, and board) at any college in the country where the foster child is accepted and, once admitted, maintains an acceptable grade point average. With some limitations imposed by Casey on the total number of students supported each year, these benefits are available to all children served by the Casey program and are considered a linchpin of the Casey system designed to encourage foster children to strive for their education goals without huge financial concerns. The Oregon and Washington public programs, in comparison, are typical of those in states terminating youth services at 18 years of age and providing no assistance for the costs of higher education.

SAMPLE

All youth who enter the Casey program do so through the public program. The public program investigates children in community settings and makes the determination, along with a juvenile court judge, whether a particular child needs a foster placement. Only after the child is in the custody of the state does Casey become involved as a placement option for youth who meet criteria stipulated by Casey. The most important criteria are that the child must be eligible for long-term placement (ie, not taken out of the home for a placement explicitly planned to be short term and not in the pool of children who are likely to be adopted), that placement must have occurred primarily because of child maltreatment (ie, not primarily because of child behavior problems or emotional problems that might require placement in a group treatment setting or in a therapeutic foster care setting), and that the child must not have a severe physical disability (eg, blindness or paralysis) or developmental disability (eg, IQ scores of less than 70) that requires special supportive services that cannot be provided in a cost-effective manner in a program as small as the Casey program. In addition, unaccompanied refugee children are not eligible for the Casey program.

Because of these selection criteria, most youth who enter the Casey system do so in late childhood or early adolescence, almost always in care for at least 1 year, and typically (in 70%-90% of cases) have a history of public program placement before entering the Casey program. The target sample for the present study was consequently selected to be youth who spent at least 12 months in foster care between the ages of 14 and 18 years during the study period (January 1, 1989, to September 30, 1998) in the Oregon or the Washington Casey or public programs. In addition, public program alumni were required to meet all other inclusion criteria for Casey placement (eg, placed primarily owing to maltreatment, with no severe physical or developmental disability) but were not placed in the Casey program because of the absence of a Casey opening at the time of their placement.

Because the number of youth in the Casey program was relatively small during the study period (+4 in Oregon and +1 in Washington), all such youth were included in the study sample. The numbers of youth in the Oregon and Washington public programs, in comparison, were many times larger. As a result, only random samples of Casey-eligible youth were selected from the Oregon and Washington public programs. These included a 4:1 sample ratio of public youth to Casey youth in Oregon (ie, 176 Oregon youth compared with 44 Casey youth) and a 3:1 sample ratio in Washington (ie, 333 Washington youth compared with 111 Casey youth). No case-level matching was used in sample selection (ie, selecting a matched set of public program alumni for each Casey alumnus). Five records of the original Oregon target sample were corrupted after data abstraction, resulting in 171 youth being included in the final target sample.
Once the random sample of public program alumni was selected, attempts were made to trace all Casey and public program alumni and perform face-to-face interviews. Alumni had been out of care from 1 to 13 years at the time of the survey, which was performed between September 2000 and January 2002. Tracing used information in case records and public records (eg, Department of Motor Vehicle records, credit records, and national mortality records). Only very small numbers of people in the target sample were found to be deceased (3 Casey alumni, 1 Oregon alumnus, and 1 Washington alumnus) or institutionalized (4 Casey alumni, 8 Oregon alumni, and 9 Washington alumni) (Figure). We were prohibited from attempting to interview institutionalized alumni by the institutional review boards at the University of Michigan and the University of Washington and of the state of Washington, so these cases were excluded from follow-up. A much more substantial loss to follow-up occurred owing to the inability to trace 29 alumni from the Casey target sample (18.7%) and comparable proportions of the Oregon (31 [18.1%]) and Washington (58 [17.4%]) target samples. Of the alumni successfully traced (76.8% of the Casey, 76.6% of the Oregon, and 79.6% of the Washington target samples), most completed the survey interview (111 Casey alumni [93.3%]; 126 Oregon alumni [96.2%]; and 242 Washington alumni [91.3%]).

INTERVIEWS

Interviews were conducted by the professional interview staff of the Survey Research Center at the University of Michigan Institute for Social Research, Ann Arbor. Each interviewer received 7 days of study-specific training and successfully completed a series of practice interviews before beginning production interviewing. The interviewers were not blinded to the agency with which the alumni had been placed because of the answers that some alumni provided regarding their foster care experience. However, other than a brief orientation to the general study aim to assess the long-term outcomes of foster youth, the interviewers were not aware that the study was concerned with comparative outcomes of Casey vs public program alumni. In addition, interviewers had no contact with any of the Casey or state study leaders during the data collection period.

The content and purpose of the survey, to assess the long-term outcomes of foster youth, were explained to potential respondents, and verbal informed consent was obtained before interviewing. These recruitment and consent procedures were approved by the institutional review boards at the University of Michigan and the University of Washington and of the state of Washington. The interview began by gathering retrospective information to supplement case records on childhood family demographics, preplacement characteristics, and foster care experiences. The remainder of the interview then focused on alumni outcomes.

Mental disorders were assessed with version 3.0 of the World Health Organization Composite International Diagnostic Interview, a fully structured lay-administered interview designed to generate research diagnoses of commonly occurring DSM-IV mental disorders. The 3 classes of disorders considered are anxiety disorders (ie, panic disorder, generalized anxiety disorder, social phobia, and posttraumatic stress disorder), major depressive episode, and substance use disorders (alcohol and other drug abuse, with or without dependence). Only disorders present in the past 12 months are considered in this report. Generally good concordance has been found between Composite International Diagnostic Interview diagnoses of these disorders and independent clinical assessments. General medical disorders were assessed with a chronic condition checklist similar to the one used in the US National Health Interview Survey. Such checklists have been shown to yield more complete and accurate reports than estimates derived from open-ended questions. Methodological studies have documented moderate-to-good concordance between such reports and medical records. The conditions considered are all those that have been shown in previous research to have meaningful prevalence in the age range of the sample and to be associated with childhood adversity: severe and/or persistent headaches, chronic back or neck pain, ulcers, arthritis, other chronic pain conditions, cardiometabolic disorders (ie, diabetes mellitus, hypertension, and heart disease), and respiratory conditions (ie, asthma, chronic obstructive pulmonary disease, and other lung diseases). Symptom-based disorders (eg, headaches) were distinguished from silent disorders (eg, hypertension) in the assessment, as respondents were asked to report whether they had each of the symptom-based conditions in the past 12 months, whether a physician had ever told them they had each of the silent conditions, and, if so, whether they continued to have these conditions in the past 12 months.

ARCHIVAL DATA ON PREPLACEMENT CHARACTERISTICS

We abstracted data from case records to obtain information about preplacement characteristics that could be used to compare survey respondents and nonrespondents and to compare background characteristics of Casey youth vs public program youth. The focus of data abstraction was on variables judged to be recorded with comparable accuracy across agencies so as not to bias comparisons. Included herein were demographics, ages at entry to and exit from foster care, reasons for placement, placement history, type and extensiveness of child maltreatment by the birth family using the Maltreatment Classification Scheme, and information about the birth family (eg, family composition, parental substance abuse, and termination of parental rights). Case record abstraction was performed by teams trained and supervised in identical ways across the different agencies to maintain comparability of coding.

ANALYTIC METHODS

The sample data were weighted using the archival data on preplacement characteristics to adjust for systematic nonresponse. We used propensity score adjustment to implement this weighting procedure. This was done by estimating a logistic regression equation separately within the Casey and public program target samples in which a dichotomous outcome variable was

<table>
<thead>
<tr>
<th>Total target sample</th>
<th>Casey Family Program</th>
<th>Oregon public program</th>
<th>Washington public program</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deceased</td>
<td>3 (1.9%)</td>
<td>1 (0.6%)</td>
<td>1 (0.3%)</td>
</tr>
<tr>
<td>Institutionalized</td>
<td>4 (2.6%)</td>
<td>8 (4.7%)</td>
<td>9 (2.7%)</td>
</tr>
<tr>
<td>Unable to trace</td>
<td>29 (18.7%)</td>
<td>31 (18.1%)</td>
<td>58 (17.4%)</td>
</tr>
<tr>
<td>Refused</td>
<td>8 (5.2%)</td>
<td>5 (2.9%)</td>
<td>23 (6.9%)</td>
</tr>
<tr>
<td>Interviewed</td>
<td>111 (71.6%)</td>
<td>126 (73.7%)</td>
<td>242 (72.7%)</td>
</tr>
</tbody>
</table>

Figure. Disposition of the target sample. Five of the original 176 case record files in the Oregon target sample were corrupted after data abstraction had ended, resulting in only 171 cases being included in the final target sample. Percentages are based on the total target sample for each program.
defined to distinguish survey respondents from nonrespondents. The predictor variables in the equations were obtained from archival intake record files about the preplacement characteristics of the respondents when they were children. We used the predicted probabilities generated from these equations to weight the data without case-level matching so that survey respondents had distributions on preplacement characteristics comparable to those of the original target sample. We then compared weighted Casey and public program samples within each state to generate a second propensity score weight that adjusted for baseline differences between Casey and public program alumni.

Multiple regression analysis was then used to estimate the significance of between-program differences in the prevalence of adult mental disorders and physical disorders among the program alumni who participated in the survey based on these doubly weighted data. We used logistic and linear link functions, depending on the outcome. Controls were included for the state (Oregon vs Washington), preplacement sociodemographics, preplacement childhood adversities, and reasons for placement. Standardized program-specific prevalence estimates were generated from the logistic regression models. Standard errors of these prevalence estimates were calculated using the jackknife repeated replications simulation method. Significance of between-program differences was evaluated at the 0.05 level with 2-sided tests.

The substantive significance of results was evaluated using the Cohen d statistic to calculate effect size. The coefficient d is the standardized (by the pooled standardized standard deviation of the outcome in the 2 samples) difference in prevalence estimates between the 2 samples. Effect sizes of 0.20 conventionally are considered small in substantive terms, while those of 0.50 are considered medium and those of 0.80 are considered large. We also calculated the number needed to treat (NNT), which is the number of youth who would have to be assigned to the Casey program rather than the public program to prevent 1 additional case of the adverse health outcome from occurring based on the assumption that mean differences reflect causal effects of the programs. The NNT is equal to the multiplicative inverse of the difference in standardized prevalence estimates of the outcome between the 2 samples. Fractional values of NNT were, by convention, rounded up to the nearest whole number.

### RESULTS

**SAMPLE CHARACTERISTICS**

Comparison of the survey respondents with nonrespondents on case record information found that a consistent substantial difference involved sex, with women consistently more likely to be respondents (61.0%-69.8%) than nonrespondents (31.0%-58.2%) (Table 1). Non-Hispanic white alumni were consistently less likely to be respondents (43.0%-58.6%) than nonrespondents (55.2%-75.6%). Respondents and nonrespondents differed much less in mean age (23.7 to 25.6 years among respondents vs 23.1 to 25.3 years among nonrespondents) and mean number of years since leaving care (5.2 to 6.4 among respondents vs 5.2 to 5.9 among nonrespondents). Some statistically significant differences were found in childhood adversities experienced by respondents and nonrespondents, especially in the Casey Washington sample, but there was little consistency in these patterns across samples. In some cases, the significant differences involved a higher prevalence of adversity among respondents than nonrespondents (eg, sexual abuse, emotional abuse, and maltreatment), whereas in other cases...
there was a higher prevalence of adversity among non-
respondents than respondents (eg, physical abuse, ne-
glect, and child behavior problems).

As noted in the “Analytic Methods” subsection of the
“Methods” section, a propensity score weight corrected
for these response biases. These weighted data were then
used to compare the Casey and public program respond-
ents to determine whether the populations treated in
the 2 programs differed significantly (Table 2). The
weighted Casey and public program samples were com-
parable in race/ethnicity (46.1% vs 45.5% non-Hispanic
white) and sex (55.4% vs 61.9% female), but the Casey
sample was somewhat older than the public sample (mean
age, 25.2 vs 23.8 years at interview) and had a some-
what shorter mean time since leaving care, although the
latter difference was confined to Oregon (4.9 vs 5.9 years).
High proportions of alumni in both systems had a his-
tory of childhood physical abuse (59.0% vs 66.3%), sexual
abuse (64.7% vs 53.2%, significantly more prevalent
among Casey than public program alumni in both states
combined), emotional abuse (89.2% vs 83.9%), and ne-
glect (58.4% vs 68.4%, significantly more prevalent among
public program than Casey alumni in Washington). Rea-
sons for placement were similar across programs, with
child maltreatment the most common reason (63.0% vs
64.0%), followed by caregiver inability to provide for the
child because of family stress or other reasons (51.5% vs
49.7%). Parental substance abuse was the only reason for
placement that differed significantly between the Casey
(46.6%) and public (21.1%) programs (χ² = 41.9, P < .001).
The propensity score weighting adjusted for these pre-
placement differences before substantive analysis.

### Table 2. Baseline Differences in Case Record Files Between Casey and Public Program Alumni

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Program</th>
<th>Oregon</th>
<th>Washington</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Casey (n=29)</td>
<td>Public (n=126)</td>
<td>Casey (n=62)</td>
<td>Public (n=242)</td>
</tr>
<tr>
<td>Sociodemographic data</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SE) age at interview, y</td>
<td>25.4 (0.4)</td>
<td>24.2 (0.2)</td>
<td>25.2 (0.2)</td>
<td>23.7 (0.1)</td>
</tr>
<tr>
<td>Mean (SE) time since leaving foster care, y</td>
<td>4.9 (0.4)</td>
<td>5.9 (0.2)</td>
<td>5.3 (0.2)</td>
<td>5.5 (0.1)</td>
</tr>
<tr>
<td>Female</td>
<td>62.7 (7.8)</td>
<td>63.3 (3.8)</td>
<td>52.8 (4.4)</td>
<td>61.1 (2.4)</td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>51.6 (8.3)</td>
<td>52.0 (3.8)</td>
<td>44.2 (4.2)</td>
<td>42.1 (2.4)</td>
</tr>
<tr>
<td>In Oregon</td>
<td>. . . . . . . . . . . .</td>
<td>. . . . . . . . . . . .</td>
<td>. . . . . . . . . . . .</td>
<td>. . . . . . . . . . . .</td>
</tr>
<tr>
<td>Maltreatment type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical abuse</td>
<td>52.9 (8.3)</td>
<td>63.0 (3.7)</td>
<td>61.2 (4.3)</td>
<td>68.0 (2.2)</td>
</tr>
<tr>
<td>Sexual abuse</td>
<td>63.5 (8.2)</td>
<td>45.4 (3.8)</td>
<td>65.1 (4.2)</td>
<td>57.2 (2.4)</td>
</tr>
<tr>
<td>Emotional abuse</td>
<td>93.1 (3.8)</td>
<td>79.9 (3.1)</td>
<td>87.9 (2.6)</td>
<td>85.9 (1.6)</td>
</tr>
<tr>
<td>Neglect</td>
<td>70.7 (8.2)</td>
<td>59.1 (3.8)</td>
<td>54.1 (4.4)</td>
<td>73.2 (2.1)</td>
</tr>
<tr>
<td>Reasons for placement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maltreatment</td>
<td>76.6 (8.1)</td>
<td>65.1 (3.7)</td>
<td>58.3 (4.3)</td>
<td>63.5 (2.3)</td>
</tr>
<tr>
<td>Parents’ substance use problems</td>
<td>58.8 (8.2)</td>
<td>17.3 (2.8)</td>
<td>42.3 (4.9)</td>
<td>23.1 (2.0)</td>
</tr>
<tr>
<td>Child behavior problems</td>
<td>12.9 (6.6)</td>
<td>27.8 (3.4)</td>
<td>16.7 (3.0)</td>
<td>18.1 (2.2)</td>
</tr>
<tr>
<td>Other</td>
<td>56.3 (8.0)</td>
<td>40.1 (3.8)</td>
<td>49.8 (4.4)</td>
<td>54.7 (2.4)</td>
</tr>
</tbody>
</table>

Abbreviations: Casey, Casey Family Program; public, public program.

*Based on propensity score weighted data that adjust for baseline differences between survey respondents and nonrespondents. Unless otherwise indicated, data are expressed as percentage (SE) of cases.

†Indicates significant difference between Casey and public program alumni at the .05 level, 2-sided test.

‡Percentage of cases in Oregon among the total population.

§Includes family stress, caregivers unable or unwilling to provide care, and other reasons.

FOSTER CARE EXPERIENCES

The hypothesis that Casey alumni had better outcomes was
based on the assumption that Casey alumni had more posi-
tive foster care experiences. This assumption was con-
firmed in an analysis of information about foster care ex-
périences abstracted from the case records (Table 3). Casey
alumni had nearly 2 more years in care than public
program alumni did in Oregon (8.4 vs 6.5; F₁,₁₅₁ = 6.6;
P = .01) and 2.5 more years in Washington (9.8 vs 7.3;
F₁,₁₃₉ = 98.4; P < .001). Casey alumni had significantly more
stable placements (mean duration of placement) than
public program alumni did in Oregon (32.7 vs 13.3 months;
F₁,₁₃₅ = 20.7; P < .001) and Washington (26.4 vs 19.2
months; F₁,₁₃₀ = 12.6; P < .001). Casey alumni were sub-
stantially less likely to experience adverse events during
their time in care. Casey alumni were at 82% (compared
with Washington alumni) to 88% (compared with Or-
egon alumni) lower risk (on the basis of risk in the public
program samples) of having experienced a reunification
failure during comparable periods of time in foster care
(2.9 incidents per 100 person-years in the Casey sample
vs 23.5 in the public program sample in Oregon
[F₁,₁₃₅ = 45.1; P < .001], and 1.9 incidents per 100 person-
years in the Casey sample vs 10.7 in the public program
sample in Washington [F₁,₁₃₀ = 68.9; P < .001]). Casey
alumni were also at a consistently lower risk of docu-
menced incidents of foster parent neglect (62.3% lower risk
in Oregon and 26.9% lower risk in Washington) and physi-
cal (81.0% lower risk in Oregon and 14.3% lower risk in
Washington) and sexual abuse (39.3% lower risk in Or-
egon and 37.9% lower risk in Washington).
CASEY VS PUBLIC PROGRAM DIFFERENCES IN ADULT MENTAL AND PHYSICAL HEALTH OUTCOMES

After adjusting for preplacement differences and differential nonresponse, Casey alumni who participated in the survey had a significantly lower 12-month prevalence than did public program alumni of all 3 classes of mental disorder studied, including major depression (11.3% vs 24.3%), anxiety disorders (28.8% vs 43.0%), and substance use disorders (5.1% vs 11.1%) (Table 4). The analysis of these effects was performed in both states combined because of the small sample sizes in each state. Effect sizes for these outcomes are all large (Cohen d=0.78-0.95), with NNTs of 8 for major depression, 7 for anxiety disorders, and 17 for substance use disorders. When combined across classes of mental disorders, Casey alumni had 44.7 fewer 12-month disorders per 100 respondents than did the public program alumni. This is equivalent to a Cohen d of 1.28 and an NNT of 3, the latter meaning that 3 youth would have had to be fostered in the Casey program rather than in the public program to prevent one 12-month case of a mood, anxiety, or a substance use disorder considered herein.

Casey alumni also had significantly lower 12-month prevalence of ulcers (7.4% vs 13.0% [Cohen d=0.65; NNT=14]) and cardiometabolic conditions (14.9% vs 22.6% [Cohen d=0.49; NNT=18]). The single disorder for which Casey alumni had a significantly higher prevalence than public program alumni involved respiratory disorders (28.8% vs 17.9% [Cohen d=0.80; NNT=10]). There were no significant differences between Casey and public program alumni, in comparison, in pain conditions (ie, chronic headaches, chronic back and/or neck pain, or other chronic pain conditions), although Casey alumni had a somewhat lower prevalence of each of these conditions than did public program alumni. When combined across all classes of physical disorders, Casey alumni had 20.1 fewer 12-month disorders per 100 respondents than did public program alumni, equivalent to a Cohen d of 0.46 and NNT of 5.

The comparatively small sample size in the study made it impossible to perform detailed analyses of subgroup variation. In light of the higher average age of the Casey alumni, however, we investigated the possibility that the significant differences between the alumni in outcomes vary by age at interview. Only 1 such interaction was statistically significant: the significantly lower prevalence of substance use disorders in the Casey compared with the public program samples, which was more pronounced (ie, the prevalence in the Casey sample became even lower relative to the public sample) with increasing age ($\chi^2=5.3; P=0.02$). Because of the wide variation in time since leaving care, we also searched for interactions involving this variable. Only 1 such significant interaction was found, involving the number of chronic physical conditions. The advantage of Casey alumni in this outcome relative to public program alumni increased significantly with increasing time since leaving care ($F_{1,421}=8.2; P=0.04$). Finally, we searched for variation in differences between the Casey and public program alumni in outcomes as a function of parental substance abuse as a reason for placement because this reason was significantly more common among Casey alumni than among public program alumni. Only 1 such interaction was found: the higher prevalence of respiratory disorders among Casey alumni was found to be confined to respondents who were placed because of parental substance abuse.

### Table 3. Foster Care Experiences of Casey and Public Program Alumni by Statea

<table>
<thead>
<tr>
<th>Experiences</th>
<th>Oregon</th>
<th></th>
<th>Washington</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Casey (n=29)</td>
<td>Public (n=126)</td>
<td>Casey (n=82)</td>
<td>Public (n=242)</td>
</tr>
<tr>
<td>Placement history, mean (SE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time in foster care, y</td>
<td>8.4 (0.3)b</td>
<td>6.5 (0.3)</td>
<td>9.8 (0.2)b</td>
<td>7.3 (0.2)</td>
</tr>
<tr>
<td>No. of placements</td>
<td>4.5 (0.4)b</td>
<td>8.2 (0.4)</td>
<td>7.3 (0.2)</td>
<td>7.8 (0.2)</td>
</tr>
<tr>
<td>Average duration of placements, mo</td>
<td>32.7 (1.0)b</td>
<td>13.3 (1.0)</td>
<td>26.4 (0.4)b</td>
<td>19.2 (0.4)</td>
</tr>
<tr>
<td>Adverse events, No. per 100 person-years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reunification failuresd</td>
<td>2.9 (1.5)b</td>
<td>23.5 (1.5)</td>
<td>1.9 (0.4)b</td>
<td>10.7 (0.4)</td>
</tr>
<tr>
<td>Incidents of neglectb</td>
<td>25.7 (7.5)b</td>
<td>68.1 (10.2)</td>
<td>52.8 (2.7)b</td>
<td>72.2 (3.0)</td>
</tr>
<tr>
<td>Incidents of physical abuseb</td>
<td>23.2 (4.7)b</td>
<td>122.3 (15.6)</td>
<td>78.2 (6.4)</td>
<td>91.3 (6.6)</td>
</tr>
<tr>
<td>Incidents of sexual abuseb</td>
<td>3.7 (0.7)</td>
<td>6.1 (0.8)</td>
<td>5.9 (0.5)</td>
<td>9.5 (0.6)</td>
</tr>
</tbody>
</table>

Abbreviations: Casey, Casey Family Program; public, public program.

a Based on propensity score weighted data that adjust for preplacement differences between Casey and public program alumni abstracted from case record files and for baseline differences between survey respondents and nonrespondents. Standardized (based on linear regression models that used preplacement characteristics obtained in case record files as the basis for standardization) means (SEs) are presented for between-system differences in outcomes.

b Indicates significant difference between Casey and public program alumni at the .05 level, 2-sided test.

c These outcomes were calculated by reweighting each respondent for the number of years individuals were in foster care and then calculating the weighted mean ratio of number of incidents per 100 person-years. This time-at-risk adjustment was used because Casey youth had longer times in care than public program youth did, leading to bias in person-level analysis without correction for time at risk.

d Defined as the number of times a child returned to his or her birth parents for 2 weeks or longer followed by a subsequent placement.

e Respondents were asked to recall the first time any of these things happened to them and were asked: “Was that a single incident, repeated (meaning the same type of abuse occurred >3 times in a 6-month period for ≥6 months), or continuous (meaning the same type of abuse occurred ≥4 times a month for ≥6 months)?” Single incidents were given a count of 1, repeated incidents were assigned a count of 18, and continuous incidents were assigned a count of 24.

### Table 4

<table>
<thead>
<tr>
<th>Experience</th>
<th>Casey (n=630)</th>
<th>Public (n=1630)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placement history, mean (SE)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time in foster care, y</td>
<td>32.7 (0.3)</td>
<td>21.6 (0.3)</td>
</tr>
<tr>
<td>No. of placements</td>
<td>4.5 (0.4)</td>
<td>8.2 (0.4)</td>
</tr>
<tr>
<td>Average duration of placements, mo</td>
<td>32.7 (1.0)</td>
<td>13.3 (1.0)</td>
</tr>
<tr>
<td>Adverse events, No. per 100 person-years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reunification failuresd</td>
<td>2.9 (1.5)</td>
<td>23.5 (1.5)</td>
</tr>
<tr>
<td>Incidents of neglectb</td>
<td>25.7 (7.5)</td>
<td>68.1 (10.2)</td>
</tr>
<tr>
<td>Incidents of physical abuseb</td>
<td>23.2 (4.7)</td>
<td>122.3 (15.6)</td>
</tr>
<tr>
<td>Incidents of sexual abuseb</td>
<td>3.7 (0.7)</td>
<td>6.1 (0.8)</td>
</tr>
</tbody>
</table>

Abbreviations: Casey, Casey Family Program; public, public program.

©2008 American Medical Association. All rights reserved.
terviewed, the interviewers were not blinded to the agencies where respondents had been placed, and the outcomes were assessed with self-reports rather than with clinical interviews and medical examinations. With regard to the first limitation, actual differences in preplacement characteristics were small. This is unsurprising because all Casey youth originally entered foster care through the public programs, were referred to the Casey program on the basis of availability of openings at the time of referral. Although propensity score weighting was used to adjust for preplacement differences, this adjustment corrected for measured preplacement variables only. Such adjustment is often no better in improving the accuracy of the estimate than the more conventional adjustment obtained using multivariate controls in conventional regression analyses. It is possible that other unmeasured preplacement differences between the 2 samples could have introduced bias into the comparisons. The second limitation (incomplete survey response) is more difficult to evaluate because it is likely that nonrespondents had worse adult outcomes than respondents did. Although similar proportions of Casey and public program alumni could not be traced and interviewed and, although preplacement predictors of response—such as family violence and parental psychopathology—were similar across samples, these were no guarantees against differential nonresponse bias. The propensity score weighting used to adjust for nonresponse bias adjusted only for differences based on observed preplacement variables, leaving open the possibility of remaining bias due to unmeasured variables that might have come into being many years after placement. The third limitation (lack of interviewer blinding) is somewhat mitigated by the fact that interviewers were not aware that the study was concerned with the comparative outcomes of Casey vs public program alumni. The fourth limitation (possible errors in outcome assessments) is somewhat mitigated by the fact that the self-reported outcome measures have been shown in previous methodological studies to be concordant with objective measures based on medical records and clinical evaluations. Because there is no reason to believe that measurement errors are systematically different for Casey vs public program alumni, measurement errors would be expected to introduce a conservative bias into the results.

The fact that the study was funded by the Casey Family Programs raises concerns about bias in the design, the data collection, and/or the data analysis. To guard against bias, the evaluation was directed by an independent investigator (R.C.K.) rather than by the Casey Family Pro-

Table 4. Adult Physical and Mental Health Outcomes of Casey and Public Program Alumni

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Casey (n=111)</th>
<th>Public (n=368)</th>
<th>Cohen d</th>
<th>NNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major depression, %</td>
<td>11.3 (0.5)e</td>
<td>24.3 (1.0)</td>
<td>0.95</td>
<td>8</td>
</tr>
<tr>
<td>Anxiety disorder, %</td>
<td>28.8 (0.9)e</td>
<td>43.0 (1.0)</td>
<td>0.91</td>
<td>7</td>
</tr>
<tr>
<td>Substance use disorder, %</td>
<td>5.1 (0.3)e</td>
<td>11.1 (0.5)</td>
<td>0.78</td>
<td>17</td>
</tr>
<tr>
<td>No. of mental disorders, %</td>
<td>65.5 (2.3)e</td>
<td>110.2 (2.3)</td>
<td>1.28</td>
<td>3</td>
</tr>
<tr>
<td>Physical disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic headaches, %</td>
<td>44.5 (1.0)</td>
<td>50.5 (0.9)</td>
<td>0.42</td>
<td>17</td>
</tr>
<tr>
<td>Chronic back or neck pain, %</td>
<td>30.0 (0.8)</td>
<td>38.1 (0.8)</td>
<td>0.66</td>
<td>13</td>
</tr>
<tr>
<td>Ulcer, %</td>
<td>7.4 (0.4)e</td>
<td>13.0 (0.6)</td>
<td>0.65</td>
<td>18</td>
</tr>
<tr>
<td>Other chronic pain conditions, %</td>
<td>24.9 (1.1)</td>
<td>27.5 (1.1)</td>
<td>0.15</td>
<td>39</td>
</tr>
<tr>
<td>Cardiometabolic conditions, %</td>
<td>14.9 (0.8)e</td>
<td>22.6 (1.1)</td>
<td>0.49</td>
<td>13</td>
</tr>
<tr>
<td>Respiratory conditions, %</td>
<td>28.8 (1.1)e</td>
<td>17.9 (0.8)</td>
<td>0.80</td>
<td>−10</td>
</tr>
<tr>
<td>No. of physical disorders</td>
<td>155.0 (2.9)</td>
<td>171.5 (2.8)</td>
<td>0.46</td>
<td>5</td>
</tr>
</tbody>
</table>

Abbreviations: Casey, Casey Family Program; NNT, number needed to treat; public, public program.

a Based on propensity score weighted data that adjust for preplacement differences between Casey and public program alumni abstracted from case record files and for baseline differences between survey respondents and nonrespondents.

b Standardized (based on linear or logistic regression models that used preplacement characteristics obtained in case record files as the basis for standardization) means (SEs) are presented for between-system differences in outcomes.

c Indicates a measure of effect size that is equal to the standardized difference in prevalence estimates across the 2 samples.

d Indicates the number of youth who would have to be fostered in the Casey program rather than 1 of the public programs to prevent 1 case of the adverse health outcome from occurring if the associations observed herein are causal. The NNT is reported only for statistically significant associations. In the case of respiratory conditions, fostering in the Casey program is predicted to increase prevalence, leading to a negative value of NNT.

e Indicates significant difference between Casey and public program alumni at the .05 level, 2-sided test.

f Includes panic disorder, generalized anxiety disorder, posttraumatic stress disorder, or social phobia. Data represent the percentage of respondents who had 1 or more of the disorders in this class.

g Includes diabetes, hypertension, or heart disease. Data represent the percentage of respondents who had 1 or more of the disorders in this class.
h Includes alcohol or other drug abuse with or without dependence. Data represent the percentage of respondents who had 1 or more of the disorders in this class.

i Estimates represent the average number of disorders per 100 people. The mental disorders outcome has a range of 0 to 7 that includes the following disorders: major depressive episode, panic disorder, generalized anxiety disorder, social phobia, posttraumatic stress disorder, alcohol abuse with or without dependence, and illicit drug abuse with or without dependence. The physical disorders outcome has a range of 0 to 11 that includes the following disorders: arthritis, asthma, chronic headaches, chronic back and/or neck pain, chronic obstructive pulmonary disease, other lung conditions, diabetes mellitus, heart disease, hypertension, ulcer, and unexplained chronic pain disorders.

j Indicates the number of youth who would have to be fostered in the Casey program rather than 1 of the public programs to prevent 1 case of the adverse health outcome from occurring if the associations observed herein are causal. The NNT is reported only for statistically significant associations. In the case of respiratory conditions, fostering in the Casey program is predicted to increase prevalence, leading to a negative value of NNT.

k Includes arthritis or unexplained chronic pain disorder. Data represent the percentage of respondents who had 1 or more of the disorders in this class.

l Indicates a measure of effect size that is equal to the standardized difference in prevalence estimates across the 2 samples.

m Indicates significant difference between Casey and public program alumni at the .05 level, 2-sided test.

n Includes panic disorder, generalized anxiety disorder, posttraumatic stress disorder, or social phobia. Data represent the percentage of respondents who had 1 or more of the disorders in this class.

o Includes alcohol or other drug abuse with or without dependence. Data represent the percentage of respondents who had 1 or more of the disorders in this class.
p Indicates a measure of effect size that is equal to the standardized difference in prevalence estimates across the 2 samples.

q Indicates significant difference between Casey and public program alumni at the .05 level, 2-sided test.
program staff. Furthermore, although the evaluation staff of the Casey Family Programs (P.J.P., J. Williams, K.O., and A.C.D.) collaborated in study design, instrumentation, and interpretation of results, research staff from the Oregon (J. White) and Washington (D.E.) public programs also collaborated in the same ways to protect against any bias in favor of the Casey program.

Within the context of these limitations, the study team found that Casey alumni had significantly better outcomes than public program alumni in all measures of mental disorder and in some, but not all, measures of physical disorder. These differences were generally large in substantive terms, with the Cohen d almost always greater than 0.5; NNTs of 7 to 17 for individual disorders, and an NTT as low as 3 in predicting the number of mental disorders. These effects are likely underestimated because the Casey program in Washington had administrative instability during the study period that led to considerably lower adherence to Casey program standards than in Oregon. Consistent with this fact, disaggregated analysis (results available on request from the authors) found less consistent evidence for better Casey program alumni outcomes than public program alumni outcomes in Washington than Oregon.

The finding of significantly more respiratory disorders among Casey alumni than public program alumni might reflect the fact that some physical disorders become more prevalent in conjunction with otherwise positive outcomes. Previous research has found, in fact, that some respiratory disorders, such as asthma and hay fever, are positively associated with social class, although most other chronic conditions are inversely related to social class. The failure to find any significant difference in the prevalence of pain conditions between Casey and public program alumni highlights the fact that the positive effects of Casey placement do not extend to all conditions.

We made no effort to introduce process measures into the prediction equations to examine pathways through which the significantly better outcomes of Casey than public program alumni were achieved. Our goal was the more limited one of evaluating whether these significantly better outcomes existed. Future analyses of the data will attempt to investigate the processes of care. Also, despite the better outcomes of Casey than public program alumni, the outcome prevalence estimates of Casey alumni are much higher than in the general population. For example, the 28.8% estimated 12-month prevalence of DSM-IV anxiety disorders among Casey alumni, although significantly lower than the 43.0% estimated prevalence among public program alumni, is still substantially higher than the 15.7% found in the subsample of people with the same age-sex composition in the nationally represented National Comorbidity Survey Replication. This means that the adverse effects of childhood maltreatment are not entirely removed by the model Casey program, even though they are significantly reduced compared with the public program.

Given the enormous societal investment in foster care, and in light of the superior outcomes of Casey alumni, it is tempting to argue on the basis of the results reported herein for widespread adoption of the Casey model in the public foster care program. This would be premature, however, because some program elements not included in the Casey program during the years studied herein, such as cognitive behavioral therapy designed specifically to treat the emotional scars of child abuse and trauma, are now available to improve on the original Casey program package. Kinship care (ie, placing the child with a responsible relative rather than with a foster family) has also become increasingly sought in the years since the study alumni were in foster care as a way to increase placement stability (a positive correlate of success). (A minority of respondents in the current study—36.1% of Casey alumni and 39.1% of public program alumni—had any kin placements during their time in foster care.) In addition, we are not sure that all of the more costly elements of the Casey package of services contribute to the greater effectiveness of the Casey program than the public programs. Nonexperimental process analysis is consequently needed to examine opportunities for refining the program.

The feasibility of implementing an expanded foster care program model in the public program, of course, depends on the willingness of state and federal legislatures to increase current funding for foster care. This willingness will be affected by many considerations. One important consideration will be the estimated cost-effectiveness of the program from a societal perspective. Calculation of such an estimate goes well beyond the scope of the current report. Further analysis of the data set is needed to evaluate cost-effectiveness, by considering outcomes in functional domains that have societal value, such as educational attainment, employment, and involvement in the criminal justice and human services systems. We plan to expand our analyses to include these domains as outcomes and to calculate the net societal benefit of expanded foster care in relation to the increased costs of a model program such as the Casey program.

Even if these analyses show that the net societal benefit of expanded foster care treatment is positive, further efforts will be needed to compare other private model foster care programs, a number of which exist around the country, with the public program in their states to assess whether the positive results found herein can be generalized to or are idiosyncratic to the time, agency practice models, and place of the present study. If the positive results can be replicated, it might be possible to perform parallel post hoc treatment process analyses to help system planners pinpoint core program components that account for these positive effects. This information could then be used to develop a blended model program for implementation in demonstration sites and subsequent dissemination to public program throughout the country if state and federal legislatures could be convinced of the importance of this undertaking. Although these results constitute only a first step in promoting this envisioned dissemination effort, they clearly document the substantial positive effects of model foster care in the domains of mental and physical health that encourage further exploration in other outcome domains.

Submitted for Publication: August 20, 2007; final revision received November 16, 2007; accepted January 5, 2008.

Author Affiliations: Department of Health Care Policy, Harvard Medical School, Boston, Massachusetts (Dr Kessler,
Mss Hiripi and Sampson, and Mr Hwang); Research Services (Drs Pecora, and O'Brien, and Downs and Mr Williams) and Systems Improvement (Dr English), Casey Family Programs, Seattle, Washington; School of Social Work (Drs Pecora and English), University of Washington, Seattle; School of Social Work, Portland State University, Portland, Oregon (Dr White); and Evans School of Public Affairs, University of Washington, Seattle (Drs Berge and Plotnick). Mr Williams is now with the Social Work Evaluation Program, University of Alaska Anchorage.

Correspondence: Ronald C. Kessler, PhD, Department of Health Care Policy, Harvard Medical School, 180 Longwood Ave, Boston, MA 02115 (kessler@hcp.med.harvard.edu).

Financial Disclosure: None reported.

Author Contributions: Dr Kessler had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Funding/Support: This study was supported by an unrestricted grant from the Casey Family Programs (an operating foundation) and in-kind support from the states of Oregon and Washington.

Additional Contributions: The staff members and agency collaborators of the Northwest Foster Care Alumni Study helped facilitate data sample selection and data extraction from case record files, the staff of the Survey Research Center at the University of Michigan carried out data collection, and the foster care alumni shared their stories and helped interpret results.

REFERENCES


