NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Costs and Effectiveness of the Fast Track Intervention for Antisocial Behavior

E. Michael Foster[*]
School of Public Health, University of North Carolina, Chapel Hill, NC, USA

Abstract

Background—Antisocial behavior is enormously costly to the youth involved, their families, victims, taxpayers and other members of society. These costs are generated by school failure, delinquency and involvement in the juvenile justice system, drug use, health services and other services. For prevention programs to be cost effective, they must reduce these costly behaviors and outcomes.

Aim—The Fast Track intervention is a 10-year, multi-component prevention program targeting antisocial behavior. The intervention identified children at school entry and provided intervention services over a 10-year period. This study examined the intervention's impact on outcomes affecting societal costs using data through late adolescence.

Methodology—The intervention is being evaluated through a multi-cohort, multi-site, multi-year randomized control trial of program participants and comparable children and youth in similar schools, and that study provides the data for these analyses. Schools within four sites (Durham, NC; Nashville, TN; Seattle, WA; and rural central Pennsylvania) were selected as high-risk based on crime and poverty statistics of the neighborhoods they served. Within each site, schools were divided into multiple sets matched for demographics (size, percentage free/reduced lunch, ethnic composition); one set within each pair was randomly assigned to the intervention and one to the control condition. Within participating schools, high-risk children were identified using a multiple-gating procedure. For each of three annual cohorts, all kindergarteners (9,594 total) in 54 schools were screened for classroom conduct problems by teachers. Those children scoring in the top 40% within cohort and site were then solicited for the next stage of screening for home behavior problems by the parents, and 91% agreed ($n = 3,274$). The teacher and parent screening scores were then standardized within site and combined into a sum score. These summed scores represented a total severity-of-risk screen score. Children were selected for inclusion into the study based on this screen score, moving from the highest score downward until desired sample sizes were reached within sites, cohorts, and conditions.

Results and Discussion—The intervention lacked both the breadth and depth of effects on costly outcomes to demonstrate cost-effectiveness or even effectiveness.

Limitations—The outcomes examined here reflect effects observed during measurement windows that are not complete for every outcome. Data are lacking on some potential outcomes, such as the use of mental health services before year 7.

Conclusion and Implications—The most intensive psychosocial intervention ever fielded did not produce meaningful and consistent effects on costly outcomes. The lack of effects through

[*]Correspondence to: E. Michael Foster, School of Public Health, University of North Carolina, Chapel Hill, Rosenau Hall, Campus Box#7445, Chapel Hill, NC 27599-7445, USA, Tel.: +1-919-966 3773, Fax: +1-919-966 0458, mike4kids@gmail.com.

high school suggests that the intervention will not become cost-effective as participants progress through adulthood.

Future Research—Future research should consider alternative approaches to prevention youth violence.

Introduction

The costs of a life of crime include government expenditures for criminal justice investigation, arrest, adjudication, and incarceration; costs to victims, such as medical costs, time missed from work, the value of stolen property as well as loss of life; and costs that accrue to the criminal and his or her family, such as lost wages. In the most comprehensive analysis of its kind, Cohen estimates that high-risk youth may generate social costs exceeding $2,000,000 in today's dollars.[1] This value is averaged across a range of likely values and based on the costs to society of three conditions: career criminal ($1.3–$1.5-million), heavy drug user ($370,000–$970,000), and high-school dropout ($243,000–$388,000).

Particularly problematic are "early starters", children whose conduct problems emerge early in life.[2] Such problems often lead to personal and social costs later in life. Those costs include (but are not limited to) criminal activities, substance use and abuse, and problems associated with early sexual debut, such as unwanted pregnancies and sexually transmitted diseases.[3]

The enormity of these costs–and the fact that a small proportion of children and youth account for a disproportionate share of crime and delinquency–suggests that society might consider devoting considerable resources to targeted prevention. Because of these high costs, even an expensive intervention may very well be cost-effective.[4] As Russell argues, however, several features of prevention work against cost-effectiveness, even for effective programs.[5] First, prevention expenditures often occur long before the resulting benefits. Because money now is worth more than money later, a dollar spent today has to generate more than a dollar's worth of benefits in the future. Second, because of imperfect targeting, prevention programs often expend resources on those who may not develop the problem of interest. Those expenditures raise the costs of the program but generate little or perhaps even no return. (Furthermore, for mental disorders, incorrectly labeling a young person as disordered can create still other costs, perhaps further offsetting any savings from serving those who actually benefit.)

These challenges are especially relevant for conduct problems. As noted, research suggests that intervention for those at risk for conduct problems should begin early in life, before a series of self-reinforcing mechanisms become entrenched.[6,7] Starting early may be necessary, but at the same time, doing so raises the bar for cost-effectiveness. Many of the costs of problem behaviors are realized during adolescence, so intervening early lengthens the time between when expenditures are made and their payoff realized.

Furthermore, other research suggests that effective interventions target multiple aspects of a child's life.[7,8] Selective interventions for high-risk youth that have shown the most promise in the short-term are those that involve multiple components, such as Tremblay's Montreal Longitudinal-Experimental Study intervention and the Coping Power Program.[9–11] While "stacking" intervention components may raise the likelihood of effectiveness, doing so raises costs as well, and the effect on cost-effectiveness is uncertain.[12]

Even though the rationale for interventions to prevent serious conduct disorder is based partly on potential cost-effectiveness, most evaluation studies have neglected the impact on

costly school services and involvement in public systems. That is, most evaluation studies have emphasized behavioral outcomes and psychiatric diagnoses rather than outcomes such as special education placements, professional mental health services received, and court adjudication. This article examines the impact of an expensive, early and sustained prevention – the Fast Track intervention – on costly outcomes and behaviors.

Fast Track is an intensive, multi-component intervention targeted to the prevention of aggression in young children. As discussed below, the intervention identified children at school entry and provided intervention services over a 10-year period. Prior outcome analyses using an intent-to-treat design suggest that over the first five years (1st through 5th grade), the intervention was moderately successful in altering developmental processes related to conduct disorder; effect sizes ranged from 2 to 5 standard deviations on child behavior problems at home and school, emergent reading skills, peer relations, and social-cognitive and emotion-coping skills.[13–16] However, during the middle school years, intervention effects were less apparent; positive intervention effects emerged on only 2 of 17 outcomes examined (children's hyperactive and self-reported delinquent behaviors).

Using data available through the end of high school, this article provides the first comprehensive assessment of intervention impact on outcomes and behaviors associated with high societal costs. If the intervention were cost-effective, one would expect to find reductions in these outcomes for the intervention group relative to the control group. We include in this evaluation all outcomes measured by Fast Track relevant to economic evaluation, including the use of health and mental health services, delinquency, involvement in the juvenile justice system, substance use, special education service use, grade retention, and school drop-out. We rely on data from a range of informants and sources, including court and medical records. Finally, this assessment employs methods described below that incorporate key study features without executing a plethora of analyses testing for various interactions. Thus, we consider a broad range of outcomes and sub-group effects while running a manageable number of statistical models. We pursue a Bayesian shrinkage strategy to avoid possible chance findings.

Method

Participants

The intervention is being evaluated through a multi-cohort, multi-site, multi-year randomized control trial of program participants and comparable children and youth in similar schools, and that study provides the data for these analyses.[13,14] Schools within four sites (Durham, NC; Nashville, TN; Seattle, WA; and rural central Pennsylvania) were selected as high-risk based on crime and poverty statistics of the neighborhoods they served. Within each site, schools were divided into multiple sets matched for demographics (size, percentage free/reduced lunch, ethnic composition); one set within each pair was randomly assigned to the intervention and one to the control condition.

Within participating schools, high-risk children were identified using a multiple-gating procedure. For each of three annual cohorts, all kindergarteners (9,594 total) in 54 schools were screened for classroom conduct problems by teachers. Those children scoring in the top 40% within cohort and site were then solicited for the next stage of screening for home behavior problems by the parents, and 91% agreed ($n = 3,274$).[17] The teacher and parent screening scores were then standardized within site and combined into a sum score. These summed scores represented a total *severity-of-risk screen score*. Children were selected for inclusion into the study based on this screen score, moving from the highest score downward until desired sample sizes were reached within sites, cohorts, and conditions. Deviations were made when a child failed to matriculate in the first grade at a core school ($n = 59$) or

refused to participate (*n* = 75), or to accommodate a rule that no child would be the only girl in an intervention group. The outcome was that 891 children (*n*'s = 445 for intervention and 446 for control) participated. Note that these levels of problems are defined relative to other children in these high-risk schools. On the kindergarten Teacher's Report Form of the Child Behavior Checklist (TRF), which provides national norms, the average Externalizing *T*-score (available for 88% of the high-risk sample) was 66.4, and 76% of these children scored in the clinical range (*T*-scores of 60 or higher.)

Across the four sites, approximately 42% of subjects are female, 50% are African-American, 41% involved subjects from a single-parent household (at the start of measurement in this study), and the overall SES was between lower and lower-middle class.

The Fast Track Intervention

The intervention was delivered in project years 2 through 11 (grades 1 through 10.)[*] During elementary school, all families were offered parent training with home visiting, academic tutoring, and social-skill training. Parent and child groups were conducted during a 2-hour "enrichment program" held at the school on Saturdays or weekday evenings. During the first 60 to 90 minutes of this enrichment program, high-risk target children met in groups of 5 or 6 in "friendship groups" led by Educational Coordinators and co-leaders, while parents met in a group led by Family Coordinators to discuss parenting strategies that would support child school adjustment and improve child behavior. In the next 30 minutes, parent-child pairs participated in positive cooperative activities and practiced positive parenting skills with staff support, called Parent-Child Sharing Time. Paraprofessional tutors worked with the children to develop reading skills in the last 30 minutes of each group meeting and twice more each week during school hours.

During year 2 group meetings were held weekly for 22 sessions, biweekly during year 3 for 14 sessions, monthly during years 4 through 6 for 9 sessions each year, and four times during year 7. In addition to the group meetings, individual support was provided to children and parents through peer pairing and home visits. Tutors provided a weekly session for pairing with non-target classmates to enhance friendships.

Children and families received a standard level of services in year 2. In subsequent years, criterion-referenced assessments were used to adjust the dosage of tutoring, home visiting, and peer pairing to match the level of functioning of each family and child. Monthly group sessions for parents and youth continued in years 6 and 7 to deal with the challenges of transitioning into middle school, resistance to drug use, and sexual development. From grades 7 through 10, individualized intervention plans were developed and implemented with each youth, based on triannual assessments of risk and protective factors. Further details on intervention components are provided elsewhere.[18]

In addition to indicated interventions, during the elementary school years, a universal social-emotional learning intervention was provided to the classrooms in which targeted youth were located. An adaptation of the PATHS Curriculum,[19] this curriculum, promoted a more competent and less aggressive social ecology and focused on the promotion of prosocial behavior, emotional understanding, self control skills, and social problem-solving skills. Elementary school teachers also received regular consultation with the Educational Coordinators, during which classroom behavior management issues were addressed.

---

[*]Since project year 1 was at kindergarten ages, year number and grade number typically differ by one. Project year number will be the same for all youth within their cohort, although grade number will differ for those who have repeated grades or dropped out of school.

Key Outcome Domains

This article examines four outcome domains: (I) health and mental health services, (II) delinquency and involvement in the juvenile justice system, (III) school failure and special education services, and (IV) substance use. Data were included given their relevance to economic evaluation as well as their availability at the time of our analyses. Table 1 provides a full listing of measures for each outcome.

Domain I: Health and Mental Health Services—Starting in year 7 of the project (grade 6 for most study children) and continuing annually through year 13, parents were interviewed using a modified, 30-minute version of the Service Assessment for Children and Adolescents (SACA).[20] Developed for epidemiological research, the SACA assesses the use of health services (primarily those related to mental health). For the preceding year, the SACA provides the parent or caregiver report of the following services: (i) total number of visits with a facility or professional on an *outpatient* basis (including mental health center, therapist, in-home family preservation worker, drug and alcohol treatment unit); (ii) total number of nights admitted to an *inpatient* facility (including psychiatric hospital, residential treatment center, emergency shelter, group home); (iii) total number of admissions to a general health provider for emotional/behavioral issues (including general hospital, emergency department, or family doctor); and (iv) total number of admissions to a general health provider for any reason (including general hospital, emergency department, or family doctor). Our assessment considered whether the youth received any services as well as the amount received.

Information on service use was also collected from the target youth in years 11–13. Information was similar to that collected from parents with the exception that general health services were not separately assessed, and outpatient and inpatient mental health services were assessed only on a yes-no basis (i.e., no counts of service visits provided). Youth also were asked about services across the past two years as opposed to within the past year. Despite these measure differences, we examine correspondence between parent and youth report of services where comparable. Finally, one item from the Life Changes assessment[21] provided an annual parent report (years 4 through 13) of whether or not the youth received medication for behavioral problems in the past year.

To supplement informant reports of health services, Fast Track implemented a review of agency records for health facilities/providers identified by the parent and/or youth. Such data provide important details on the use of health services such as costs and types of services delivered. These records provide information for which parents were not asked (e.g., service costs) as well as information that can be used to assess the reliability of caregiver report (e.g., days of service use). For this article, we use outcomes from the records review to provide alternative measures of services use. Two outcomes – costs and number of services received – were selected given their relevance to economic evaluation. Appendix A describes the steps involved in collecting and processing medical records.

Domain II: Delinquency and Involvement in the Justice System—Data for this domain were obtained from parent and youth SACAs, court records, and youth self-report of delinquency. Both parent and youth SACAs also measured youth involvement in the juvenile and criminal justice systems. For this domain, we considered the following measures: (i) number of police contacts (not involving arrest), (ii) number of arrests, (iii) number of court appearances, (iv) total number of nights incarcerated in a detention center or jail (including time served before or after trial), and (v) whether or not a guilty judgment was rendered regarding a crime committed in the past year (the latter from the youth SACA). These outcomes were examined from years 9 through 13.

Beginning in year 7, court records were collected annually from the jurisdiction where the youth lived (or surrounding counties if applicable). Permission to search juvenile court records was required for youth outcomes as well as administrative orders from local jurisdictions. We examined outcomes from the court records through year 13, focusing on three key constructs: (i) number of reported crimes for which the youth was arrested (excluding status offenses), (ii) number of severe crimes (designated as crimes where a person is harmed or the potential for harm to persons is high), and (iii) whether or not the youth had been jailed. We also examined the impact of the intervention on age at first arrest for any crime and for severe crimes only.

The Self-Report of Delinquency (SRD)[22] was administered to youth starting in year 8 and included here through year 11. Our analyses used the antisocial behavior subscale derived by Fast Track investigators as a summary score of the amount of delinquent behavior that was self-reported (this summary scale excluded less-serious status offense items but included more serious acts such as "stolen an item greater than a hundred dollars in value," "sold heroin or LSD," "attacked someone with intent to hurt," and "had sex with someone against their will"). A frequency score was created by summing across the 25 scale items measuring the number of behaviors that occurred, and then transformed to remove skew using a square-root transformation.

Domain III: School Failure and School Services—School outcomes were obtained from school administrive records.[23] For our report we focused on three key outcomes: (i) special education services, (ii) whether the youth had repeated a grade, and (iii) whether the youth graduated from high school. For the first two outcomes, data were structured so that we could examine the likelihood of these outcomes having occurred (in any grade from 1st grade through high school graduation age) as well as the time until such school events first occurred.

Domain IV: Substance Abuse—These analyses were based on two measures. First, using questions from the *National Longitudinal Study of Adolescent Health*,[24] youth self-reports of tobacco use and alcohol/drug use were collected from years 8 through 13. The analyses considered three outcomes: (i) number of days used cigarettes (past month), (ii) number of days been drunk on alcohol (past year), and (iii) number of times used marijuana (past month). Parents reported youth involvement with substance use in years 9 through 13 on the Parent Daily Report.[25]

Analytical Model

Pooling across Sub-groups—Potential heterogeneity of the treated and moderation of effects shaped our analytic approach. Moreover, we wanted to establish models that would recognize that intervention impact might vary across key sub-groups: project site, gender, and initial risk status. Regarding the last, previous research has shown differences in the outcome of adolescent externalizing behavior disorders; intervention effects were greater for children showing higher risk screen-scores at baseline.[26] However, obtaining separate estimates for the various sub-groups defined by site, gender and risk-status would lead to numerous underpowered analyses and to potential chance findings resulting from multiple testing.

For that reason, we employed Bayesian subset analysis.[27–29] This approach is a form of empirical Bayesian estimation; it incorporates the prior beliefs of the intervention developers that the intervention would work for all sub-groups. (One can see this belief, for example, in the power calculations used to design the study.) The analysis incorporates this belief rather than fully imposing it and allows the data (through the likelihood function) to push the

estimates toward sub-group specific estimates. The final estimates are a compromise between a single one-estimate-fits-all estimate and sub-group-specific estimates; the variation across sub-groups are "shrunk" toward the prior belief (no variation across subgroups). The degree of the shrinkage can vary across outcomes and sub-groups depending on a range of factors, such as the size of the sub-groups.[30] This approach is much like one might use in meta-analysis of multiple studies or multi-level modeling of neighborhood effects. One might use the shrinkage estimators to characterize a given neighborhood's random effect in light of the prior belief that the processes of interest were the same across neighborhoods.[31] The specific statistical test one might use to test for effect heterogeneity is the same as in meta analyses and are described below.

K aspects of the study design–site, gender and initial risk status (4 X 2 X 2)—defined sixteen subgroups. For outcomes with low occurrence we used an 8-subgroup designation (4 X 2) based on site-by-risk status. This adjustment was especially necessary for delinquency outcomes such as arrest where the number of female offenders was very low. (Those models included gender as a covariate in outcome models.) This approach involved a test of the null hypothesis that intervention effects are uniform across the subgroups (the Q-test). As in meta-analysis, lack of uniformity (as indicated by a significant result on this test) implied that we use should random-effects estimates that recognize a higher level of variation across sub-groups. Alternatively, we used fixed-effects estimation to pool the estimates if this test indicated a nonsignificant result [32, 33]. When we examined subgroup-specific effects, we focused on the empirical Bayes estimates. Those effects are a weighted average of the overall effect and the effect for the subgroup. Such estimates have superior statistical properties.[33] Regardless of the results of study heterogeneity tests, this approach further allowed us to evaluate subgroup effects, especially as far as certain key patterns might be apparent (e.g., effects occur for a subgroup across a range of outcomes).

Modeling Outcomes—Regression models were selected based on the distribution of the outcomes. Given that many of the outcomes involved counts (e.g., number of arrests), negative binomial models were used given their appropriateness for the highly dispersed count variables seen in this study.[34] For models assessing binary outcomes –often whether or not an event occurred—we used logistic regression. In the few cases where outcomes were normally distributed, we used regular regression. In the presence of extreme values/outliers, we used either a log transformation or square-root transformation before conducting analyses.

Some outcomes involved the time until an event or outcome first occurred, such as arrest for a serious offense. We employed the Cox proportional hazard model for the outcomes where an exact date was available for event occurrence (i.e., arrest date) and discrete hazard models when event occurrence had a wider interval (such as first year receiving special education services).

Participant Characteristics at Baseline—At the beginning of the study, the Fast Track principal investigators identified 25 covariates that represent key characteristics of the participating subjects/families, and these have been used consistently in various analyses of intervention impact.[26] These variables are listed and described in more detail in Appendix B. In hindsight, this list seemed rather short of demographic characteristics. For these analyses, we added four background variables representing parent and household characteristics: whether the biological father was in the household, whether the mother had been a teenager at first childbirth, maternal education, number of children in the household. Collectively, when analyzing study outcomes, these variables controlled for pre-existing intervention group differences.

Because sample sizes for some subgroups were small, including all of the covariates as regressors was impractical. For that reason, we ran a logit model with intervention status as the dependent variable and the covariates as explanatory variables. We then calculated the predicted probability of intervention status, the propensity score. As demonstrated by Rubin and Rosenbaum, the propensity score is a combination of the covariates where the weighting reflects the covariates' potential for confounding.[35]

Missing Data—By year 13 of the project, missing data rates for the entire sample reached 25% to 30% (these rates vary across measures and data sources). For the SACA, 30% of the original sample had missing data by year 13. This rate did not differ significantly by intervention status: the rates were 35% and 31% for the control group and intervention group, respectively. Missing data rates for the court records were slightly lower than that for general informant report measures, averaging a little over 11% among control and intervention groups for the available data. In order to accommodate missing data appropriately, we employed multiple imputation (MI) as implemented in IVEware.[36] This program provides data imputation when the data include outcomes that are unordered categorical, ordered categorical, count, and continuous. IVEware methods aassumes data are missing at random (MAR), and we assume that our models satisfy this assumption given the inclusion of multiple measures of baseline and follow-up participant characteristics. All outcomes used in the primary analyses were included in the MI models, and analyses were based on five imputed datasets.

MI estimation was performed separately by intervention status. Doing so maintained important interactions between intervention status and other key constructs. (Thus, interaction terms did not need to be included in the MI models). Separate imputation models are important if patterns of missingness differ between groups. We did not impute separately by site because of sample size. In that sense, our imputation model is more restrictive than our analytical model.

The imputation model included the 29 covariates described above as well as 4 school-level characteristics (student-teacher classroom ratio, percentage of white students in school, percent free-lunch eligible, percentage male students) and 3 district-level characteristics (per-capita income, median home value, poverty rate) all of which might influence the likelihood of missing data. All school- and district-level variables were averages child across years.

Results

Table 2 provides intervention effects across the outcomes described above, including pooled estimates. The table lists fixed-effects estimates if study heterogeneity tests were significant and random-effects estimates if heterogeneity tests were not significant at p<.05. Because the underlying statistical models differed across outcomes, the coefficients in Table 2 (column 4) have different interpretations and are not comparable across models. For example, for a dichotomous outcome Table 2 presents the beta estimate from logistic regression; for a continuous outcome, the beta is the ordinary regression coefficient (or adjusted mean difference). These estimates are not directly comparable, and using them to assess whether the intervention had bigger effects for some outcomes than for others is difficult. The estimates, however, were directly estimated by the statistical procedures and are presented for that reason.

To facilitate comparison of intervention effects across outcomes, Table 3 provides marginal effects. These represent the effect of the explanatory variable on the outcome variable in its natural metric (holding all other model covariates constant.[37] For example, in logistic

regression the marginal effect represents the effect of the intervention on the predicted probability of the event. (Given that we could not derive marginal effect estimates through the meta-analytic routines in Stata, these effects were approximated using regression models run on the multiply imputed datasets with sub-groups (site, gender, and risk status) included as covariates.)

Domain I: Health and Mental Health Services

Using parent reports of health and mental health services (SACA) for years 7 through 13, we examined service use. Intervention youth were less likely to use general health services for any reason ($p<.05$) or to use general health services for mental health problems ($p<.01$). Control parents indicated a significantly higher number of general health services received for any reason ($p<.01$); no significant difference was detected for the amount of general health care for mental health purposes. No effects were apparent for inpatient and outpatient mental health services. When combining all 7 years, intervention parents did report greater use of inpatient mental health facilities ($p<.10$).

We also examined youth reports of service use. Despite differences in measurement (described below), we anticipated rough agreement between parental and youth reports of mental health service use (general health services use was not measured in the youth SACA). Unlike parental reports, youth reports revealed no intervention effect on inpatient services, and youth in the control condition were more likely to report having received outpatient services for mental health problems ($p<.01$).

To better compare between parental and youth reports of services, we limited parent reports to the same period for which youth reports are available (years 10–13). This narrowed range produced better agreement between sources. For outpatient mental health services, intervention parents indicated lower likelihood of service receipt ($p<.05$) and fewer services received ($p<.01$). No significant intervention effects were detected in parent report of inpatient mental health service use or amount.

Data on the use of medication for emotional or behavioral problems (parent report) was available for 12 consecutive years (covering years 2 through 13). No intervention effects were apparent.

Where differences are statistically significant, how large are they? Table 3 provides marginal effects for services use. Parents indicated that intervention youth were almost 8% less likely to seek services for mental health assistance from a general health provider and 16% less likely to seek services from a general health provider for any reason. Youth indicated that the probability of receiving outpatient mental health services was 24% lower among intervention participants; that same effect was 14% lower based on parent reports.

Table 3 also provides the marginal effects for the amounts of service use. Specifically, intervention youth received 1.14 fewer general health services (considering years 7 through 13), and 64 fewer outpatient mental health services (considering years 10 through 13). In general, intervention effects were small.

We also examined intervention effects on service expenditures using billing records. Our analyses focused on the total amounts spent across all years and service sectors. Cost data were challenging to analyze given their skewed nature. We used two approaches: First, we created cost categories for years 9 through 13. (The categories were quintiles or tertiles depending on the variability in costs.) We then used ordinal logit models to assess intervention effects on costs incurred. We also ran negative binomial models on a square-root transformed version of the total costs outcomes. As with the primary outcome models,

we used multiply-imputed agency data, and models included controls for site, gender, screening risk level, and propensity for treatment. Model standard errors were adjusted to reflect the clustering of participants in agencies. (See Appendix A for more details.)

We first examined expenditures regardless of whether they are related to emotional and behavioral problems per se. No intervention effect was found the amount (negative binomial models) or level (ordinal logit) of cost. Considering agency records regardless of service sector, we also examined the number of services received by sector. Separate models were estimated for services including general health, mental health, pregnancy-related (females only), or medication-related. None of the models revealed an intervention effect.

We next assessed sector-specific total costs. The logit models and negative binomial models were used to examine intervention effects on inpatient mental health costs, outpatient mental health costs, general health costs, and general health costs only used for mental health purposes. For outpatient mental health, the two techniques produced discordant results. Ordered logit indicated no intervention effect, but intervention youth incurred lower outpatient mental health costs across in the alternative model ($p<.05$).

Medical records and parental reports of service use for emotional and behavioral problems involving general health providers agreed. For both the logit model ($p<.01$) and the model of transformed total costs ($p<.01$), intervention youth incurred lower costs. Medical records and parental report for general health services for any reason did not agree: In this case, the medical records revealed no intervention effect. Nor did medical records reveal any difference in inpatient mental health service use.

Domain II: Delinquency Cutcomes

Parent and Youth Report—Using SACA data, we examined four delinquency outcomes: police contact, arrests, court appearances, and detention center admissions. We first examined whether these outcomes occurred, based on separate statistical models for parent and youth data. No intervention effects were found for any of the outcomes for either informant. Next, we ran models examining the levels of these four outcomes. Based on youth report, intervention participants indicated fewer arrests over years 11 through 13 than control youth ($p<.01$). This same outcome reported by the parents – covering years 9 through 13 – indicated no intervention effect. Limiting the parental reports to the same years did not reconcile these reports. Finally, analyses of an additional item in the youth SACA – whether found guilty of a crime in the past year –again revealed no intervention effect.

SRD data provided information on antisocial behavior were available for years 8 through 13. Negative binomial models indicated no effect on the level of antisocial behavior, but control youth were more likely to report any antisocial behavior ($p<.01$).

Table 3 presents the corresponding marginal effects. Intervention youth averaged .75 fewer arrests in a given year. Intervention youth were 20% less likely to exhibit any antisocial behavior.

Court Records—Analyses of these data focused on three outcomes. We examined the likelihood of and number of "severe" crimes committed. Also examined were the likelihood and amount of all crimes (including less severe crimes). Intervention effects were found for neither an indicator of any offense or number of offenses for severe or all crimes. The analyses also considered whether the youth had ever spent time in jail (through year 13). No intervention effect was apparent.

Finally, using court record data, we examined age at first criminal activity using survival analysis, considering severe and any crime committed separately. No intervention effect was apparent in either analysis

Justice System Data From Facility Records—The project obtained records from juvenile justice facilities. We assessed costs related to justice system services as well as the overall count of such services. No intervention effects were apparent for expenditures or for the level of juvenile justice involvement.

Domain III: School Failure and School Services

School data were available through year 13. We examined intervention effects for three outcomes-use of special education, grade retention and high-school completion. No intervention effects were apparent. We also considered the timing of school services. We used discrete-time hazard models to examine time to first retention or to entry into special education. No intervention effects were apparent.

Domain IV: Substance abuse

Table 2 shows the results of the pooled estimates for the substance abuse outcomes. No intervention effects were apparent for neither parental reports of substance abuse by the youth nor youth reports of tobacco, alcohol (drunkenness), or marijuana use (Tobacco, Alcohol and Drugs).

Variation Across Subgroups

The right column in Table 2 presents significant study-level effects based on empirical Bayes estimates. (Initials indicate the subgroup involved if significant; effect sizes are shown in parentheses). These separate tests allowed us to examine patterns of larger or smaller effects fo sub-groups across outcomes.

In general, we find no consistent patterns among subgroups, and many sub-group effects are in the wrong direction. Based on the lack of consistent patterns, we did not assess potential sub-group differences across these outcomes any further.

Net Benefits Associated with Intervention Effects

According to prior analyses, the intervention cost approximately $58,000 per child.[4] (Costs were estimated from a payor perspective).[38] The key question is whether reductions in costly services and outcomes might offset the costs of the intervention.

As noted above, no clear pattern of intervention effects emerged from our analyses, and for that reason, we do not present a full economic evaluation. From the perspective of the intervention, the most promising effects involve the use of health and mental health services. We determined whether expenditures on these services might have offset intervention costs. Using medical records, average costs per youth were calculated for services revealing an intervention effect: outpatient mental health costs (for the project years 10–13), general health provider for family-indicated mental health purposes, and general health providers for general health purposes.

For outpatient mental health services, control group participants averaged $6,249 on average across the years assessed. The figure for the intervention group was $4,905, representing a savings of $1,344. For general health services (delivered for any reason), the control and intervention groups totaled $6,572 and $5,466, respectively. These differences represent a further savings of $1,106. Taken together, while these savings represent a significant portion

of overall expenditures on those services, they offset relatively little of the costs of the intervention.

Discussion

This article provides the first and only comprehensive assessment of the Fast Track intervention on outcomes most relevant to an assessment of program costs and benefits. We examined these outcomes fully incorporating key sample characteristics (such as initial risk level) while avoiding ad hoc exploration of the data. The findings reveal lower use of and cost for health and mental health services by intervention participants. Based on youth report, we find lower levels of antisocial behavior and justice system involvement among intervention participants, but parent reports did not confirm these differences. No intervention effects were evident in other domains, such as substance use, special education services, school drop-out, or grade retention. We also find no consistent evidence that a sub-group exists for which the intervention was effective and to which a future, cost-effective implementation might be targeted.

In sum, these analyses suggest that the intervention lacked both the breadth and depth of effects on costly outcomes to demonstrate cost-effectiveness or even effectiveness

Future analyses may reveal nuanced effects for sub-groups or for fine-grained outcome measures. However, those analyses will have to be interpreted with caution, especially if they involve sub-groups or outcomes that are not anticipated in advance. A project of this size is often subjected to many, many analyses, and without careful attention, chance findings will appear and be interpreted as real insights. By the standards of research on clinical trials, such findings should be considered exploratory at best.[39]

Assessing Cost-Effectiveness

This article does not include a full calculation of net health benefits or other global assessment of the return on investment.[40–42] The lack of intervention effects made such a calculation superfluous. The best foundation for a positive return on investment involved health services, and as we discuss above, these savings offset only a small portion of the interventions costs.

However, one option would be to calculate net benefits using all outcomes, even those for which the intervention effect was not statistically significant. While some would argue that statistical significance is irrelevant for economic analysis,[43] insignificant effects typically reflect both small effects and large confidence intervals. In the case of the former, the impact on the economic bottom line is typically small unless the behavior involved is very costly to society. Interpreting such effects is a risky business. As often as not, insignificant effects may be of unanticipated direction; it is not a given, therefore, that including such effects will improve the returns to the intervention. Whether a tally of the many null effects would improve even a rough estimate of net benefits is unclear.

Insignificant outcome-specific effects often reflect large confidence intervals, and the imprecision involved has important implications for the economic analysis. An essential element of a net benefits calculation is a confidence interval (or a cost-effectiveness acceptability curve).[12,44,45] The uncertainty surrounding net benefits reflects the uncertainty surrounding each intervention effect included. One can see, therefore, that adding insignificant findings to the calculation of net benefits will inflate the resulting confidence interval.

Could the Intervention Become Cost-Effective Over Time?

Could future analyses reveal that the intervention is cost-effective? Such a development could come from effects appearing where they do not exist now. The marginal effect on likelihood for jail would provide more promise for cost-effectiveness given the potential for future economic benefits related to such an outcome. However, given what the intervention costs, an effect on time in jail of the size we find here (a 13.5% reduction) is unlikely by itself to offset program costs for the whole sample. One would have to discount any savings back to the first year of the intervention. Even at a modest discount rate (5%), the effect of discounting is substantial: A dollar saved in year 13 of the study represents only 53 cents at the start of the study. It seems likely that policy maker's rate of time preference is even larger – at 10%, a dollar saved is worth only 29 cents. Aos *et al.* estimate that the costs of incarceration are roughly $30,000 per year.[46] In that case, the effect estimated here would represent $1,500 for each year persisted for the average program participant. Obviously, such an effect would have to persist for decades to justify such a costly intervention.

Limitations

First, the outcomes examined here reflect effects observed during measurement windows that are not complete for every outcome. Data are lacking on some potential outcomes, such as the use of mental health services before year 7 (when the SACA was added to the study). In that light, the estimates provided here are conservative. It is worth noting that the coverage of key outcomes is best during adolescence, when many of the outcomes are most likely and most costly.

Other limitations involve broader issues reflecting the overall study design and its execution. Even though the sample size is large, randomization involved nine pairs of schools. As a result, the results are potentially subject to unobserved differences between groups. With only nine units, the power of randomization to balance unobserved factors is low. However, the ability of such unobservables to bias findings seems rather limited. After all, these factors would have to involve school-level differences, the effect of which persisted for years after the children left those schools. Such lasting differences seem somewhat unlikely, given that original schools were drawn from the same communities.

A third limitation also involves study design: while our analyses incorporate site-level differences, the study includes only four sites. Four sites are enough to identify between-community differences but insufficient to really unpack those differences into contextual factors, such as juvenile justice policies or district-level factors. Future research might consider more sites. While daunting in terms of scope and expense, such research has been conducted in other areas of human services, such as job training or welfare policy. That research indicates substantial variability in program impact across communities differing in local labor market conditions, economic factors, and so on. However, such differences also make it difficult to do anything other than speculate how a modified intervention might perform in a new community where these conditions differed.

Broader Implications for Prevention Research

When any intervention fails to produce anticipated effects, one must return to the original program model[47]. In particular, developmentalists and prevention scientists need to consider whether and how the underlying theories involving social cognition and other psychological processes really influence individual behavior (e.g., that aggression *causes* poor school performance). Much of the research linking social cognition and other psychological measures and children's behavior relies on associations. It remains an open question whether these relationships are truly causal or whether social cognition and aggression both reflect other, unmeasured influences, influences apparently not affected by an intervention like Fast

Track. Claims that "individual behavior is just hard to change" are not likely to lead to real improvements in the theories underlying the intervention. Researchers need to look seriously at the underlying scientific foundations of a generation of interventions. Fast Track relied on the best developmental theory had to offer, and it was largely ineffectual.

The lack of findings raises other questions about assumptions commonly shared in prevention research. As noted, perhaps better targeting might improve the cost-effectiveness of the intervention. One barrier to the cost-effectiveness of intervention in any area of health or mental health is the fact that prevention often involves devoting resources to subjects who may not develop the disorder of interest.[5] In that case, waiting until children are older to start the intervention might improve cost-effectiveness: as children demonstrate problem behavior for a longer period, it may become more apparent which children really need intervention. Early intervention may improve effectiveness, but early may not be the most cost-effective. Relatedly, schools may not be the best setting for identifying children needing intensive intervention. One might look to other social systems, such as child welfare, to identify those children most needing intensive and costly intervention.

The evaluation itself raises other important questions. As became apparent through reanalysis of who is in the study, the recruited sample was more diverse than originally reported in multiple publications. The intervention involved not just the most poorly behaved 10% of children screened as claimed for a decade. The argument that study children were too severely challenged does not explain the small and inconsistent effects of the intervention. A related issue involves the families who would not participate – who are these children and families who were more difficult to reach? Perhaps the intervention would have proven cost-effective for these families. Tracking and recruiting these families at additional expense might have proven cost-effective, even if it meant reducing the resources devoted to the intervention itself. At this point, we do not know. Very little information is available on those children and families.

In many ways, the intervention reflects the best of prevention research but also its weaknesses. For example, many developmental studies are based on samples that are not representative (e.g., the NICHD Daycare study) of the US population or even any population that one might describe easily. That the evaluation was wrong about who was in the study hardly distinguishes it within developmental psychology.

Another weakness of developmental psychology that shaped the Fast Track project is the lack of data sharing. For a decade, the investigators did not share data on the intervention. A fuller and timelier assessment of the intervention's impact may very well have emerged were data sharing the norm. Program developers naturally face a conflict of interest in assessing their own interventions.[48] Data sharing would allow more researchers to examine the data and to understand the processes at work. Even if these researchers replicated the findings of the program's developers, such confirmation is reassuring. And with data so complex, additional analyses are likely to shed new light on whether and how the intervention worked. A handful of researchers can analyze multiple waves of data from multiple sites with multiple informants covering multiple domains only so fast. Data sharing would have allowed for a fuller assessment of the intervention in a timely manner.

In addition, the history of the project reflects the fixation of developmentalists on their theories, but at times theory is the proverbial tail that wags the dog. Publications from the project on issues of theory outnumber those on the intervention by a ratio of at least five to one. In an intervention project, the bottom line on effectiveness has to be the real bottom line in terms of focus of the research enterprise. That does not appear to have been the case

with Fast Track. This imbalance of resources explains much of the delay in reporting evaluation outcomes.

Other weaknesses in developmental psychology and prevention research also are apparent in the study. Many analyses in the field follow a meandering analysis plan—researchers run analyses, change the model, and then run more analyses, increasing the likelihood of chance findings. (For example, many papers in psychology rely on modification indices in structural equations software in spite of numerous cautions about their poor statistical properties.) In contrast, the standards of clinical trials specify a predetermined analysis plan, and those standards should guide the evaluation of preventive interventions. A linchpin of those trials is that "the extent to which ... the primary analysis is planned a priori will contribute to the degree of confidence in the final results and conclusions of the trial", and "only results from analyses envisaged in the protocol (including amendments) can be regarded as confirmatory".[39]

In many instances, researchers undertake these nuanced analyses to gain insights into developmental theory. But the fact of the matter is that developmental theory is so non-specific that it can be used to explain any finding post-hoc. The reality is that analyses can proceed until some chance finding proves sufficiently interesting, and theory is then applied. Furthermore, many of the theory-informed analyses involve complicated statistical issues, and developmental psychology has often gotten them wrong. For example, many developmentalists are interested in whether some psychological mechanism mediates the effect of their intervention. However, the construct involved is often a "collider", and the resulting estimates do not describe direct or indirect effects.[49] Developmental theory also can suggest moderation, and those analyses are often problematic as well. The putative moderator may involve a self-selected state, and in that case, the differential "effect" of the intervention is really a case of differential self-selection in the moderating condition. The bottom line is that even in a randomized trial, including mediators or moderators that are not randomized raises difficult statistical issues, whether the addition of those variables is informed by developmental theory or not.

Analysis like the ones presented here are broad in nature, but they could have been planned prior to the study. Such analyses should be reported annually and would provide a background for assessing the more nuanced and in-depth assessments that emerge more slowly. Analyses like these do not need to be the last word in an evaluation, but they do need to be the first (and timely) word.

Acknowledgments

References

1. Cohen MA. The Monetary Value of Saving a High-Risk Youth. J Quant Criminol. 1998; 14:5–33.

2. Moffitt TE. Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. Psychol Rev. 1993; 100(4):674–701. [PubMed: 8255953]

3. Capaldi, DM.; Kim, HK.; Short, JW. Woman's involvement in young adult romantic relationships: A Developmental-contextual model. In: Bierman, K.; Puttalez, M., editors. Aggression, antisocial behavior and violence among girls: A Developmental perspective. New York: Guilford Publications Inc; 2003.

4. Foster EM, Jones D. Conduct Problems Prevention Research Group. Can a costly intervention be cost-effective?: An analysis of violence prevention. Arch Gen Psychiatry. 2006; 63:1284–1291. [PubMed: 17088509]

5. Russell, S. Is Prevention Better than the Cure?. Washington, DC: The Brookings Institution; 1986.

6. Webster-Stratton C, Taylor T. Nipping early risk factors in the bud: Preventing substance abuse, delinquency, and violence in adolescence through interventions targeted at young children (0–8 years). Prev Sci. 2001; 2(3):165–192. [PubMed: 11678292]

7. Conduct Problems Prevention Research Group. A developmental and clinical model for the prevention of conduct disorders: The Fast Track Program. Dev Psychopathol. 1992; 4:509–527.

8. Henggeler, SW.; Schoenwald, SK.; Borduin, CM.; Rowland, MD.; Cunningham, PB. Multisystemic Treatment of Antisocial Behavior in Children and Adolescents. New York: The Guilford Press; 1998.

9. Lochman JE, Wells KC. Effectiveness of the Coping Power Program and of classroom intervention with aggressive children: Outcomes at a 1-year follow-up. Behav Ther. 2003; 34:493–516.

10. Pepler DJ, King G, Craig W, Byrd W, Bream L. The development and evaluation of a multi-system social skills group training program for aggressive children. Child Youth Care Forum. 1995; 24:297–313.

11. Tremblay RE, Pagani-Kurtz L, Masse LC, Vitaro F, Pihl RO. A bi-modal preventive intervention for disruptive kindergarten boys: Its impact through mid-adolescence. J Consult Clin Psychol. 1995; 63:560–568. [PubMed: 7673533]

12. Foster EM, Olchowski AE, Webster-Stratton CH. Is stacking intervention components cost-effective? An analysis of the Incredible Years program. J Am Acad Child Adolesc Psychiatry. 2007; 46(11):1414–1424. [PubMed: 18049291]

13. Conduct Problems Prevention Research Group. Initial impact of the fast track prevention trial for conduct problems: II classroom effects. J Cosult Clin Psychol. 1999; 67(5):648–657.

14. Conduct Problems Prevention Research Group. Initial impact of the fast track prevention trial for conduct problems: I The High-risk sample. J Consul Clin Psychol. 1999; 67(5):631–647.

15. Conduct Problems Prevention Research Group. Evaluation of the first three years of the Fast Track Prevention Trial with children at high risk for adolescent conduct problems. J Abnorm Child Psychol. 2002; 30:19–35. [PubMed: 11930969]

16. Conduct Problems Prevention Research Group. The effects of the Fast Track program on serious problem outcomes at the end of elementary school. J Clinical Child Adolesc Psychol. 2004; 33:650–661. [PubMed: 15498733]

17. Lochman JE. Conduct Problems Prevention Research G. Screening of child behavior problems for prevention programs at school entry. J Consult Clin Psychol. 1995; 63(4):549–559. [PubMed: 7673532]

18. Bierman KL, Cole JD, Dodge KA, Greenberg MT, Lochman JE, McMahon RJ. Implementing a Comprehensive Program for the Prevention of Conduct Problems in Rural Communities: The Fast Track Experience. Am J Community Psychol. 1997; 25(4):493–514. [PubMed: 9338956]

19. Kusche, C.; Greenberg, M. The PATHS Curriculum. Seattle, WA: Developmental Research and Programs; 1993.

20. Stiffman AR, Horwitz SM, Hoagwood K, et al. The Service Assessment for Children and Adoldescents (SACA): Adult and child reports. J Am Acad Child Adolesc Psychiatry. 2000; 39(8):1032–1039. [PubMed: 10939232]

21. Dodge KA, Bates JE, Pettit GS. Mechanisms in the cycle of violence. Science. 1990; 250(4988): 1678. [PubMed: 2270481]

22. Elliott, DS.; Huizinga, D.; Ageton, SS. Explaining delinquency and drug use. Beverly Hills, Calif: Sage Publications; 1985.

23. Walker, HM.; Block-Pedego, A.; Todis, B.; Severson, H. School Archival Records Search (SARS): User's guide and technical manual. Longmont, CO: Sopris West; 1991.

24. Resnick MD, Bearman PS, Blum RW, Bauman KE, Harris KM, Jones J, Tabor J, Beuhring T, Sieving RE, Shew M, Ireland M, Bearinger LH, Udry JR. Protecting adolescents from harm. Findings from the National Longitudinal Study on Adolescent Health. JAMA. 1997; 278(10):823–832. [PubMed: 9293990]

25. Chamberlain P, Reid JB. Parent observation and report of child symptoms. Behav Assess. 1987; 9:97–109.

26. Conduct Problems Prevention Research Group. Fast Track Randomized Controlled Trial to Prevent Externalizing Psychiatric Disorders: Findings From Grades 3 to 9. J Am Acad Child Adolesc Psychiatry. 2007; 46(10):1250–1262. [PubMed: 17885566]

27. Dixon DO, Simon R. Bayesian subset analysis. Biometrics. Sep; 1991 47(3):871–881. [PubMed: 1742443]

28. Dixon DO, Simon R. Bayesian subset analysis in a colorectal cancer clinical trial. Stat Med. 1992; 11(1):13–22. [PubMed: 1557569]

29. Simon R. Bayesian subset analysis: application to studying treatment-by-gender interactions. Stat Med. 2002; 21(19):2909–2916. [PubMed: 12325107]

30. Carlin, JB.; Louis, TA. Bayes and empirical Bayes methods for data analysis. 2. New York: Chapman and Hall/CRC; 2000.

31. Raudenbush, SW.; Bryk, AS. Hierarchical linear models: applications and data analysis methods. 2. Thousand Oaks: Sage Publications; 2002.

32. Lipsey, MW.; Wilson, DB. Practical meta-analysis. Thousand Oaks, Calif: Sage Publications; 2001.

33. Petitti, DB. Meta-analysis, decision analysis, and cost-effectiveness analysis: methods for quantitative synthesis in medicine. 2. New York: Oxford University Press; 2000.

34. StataCorp. Stata Statistical Software: Release 10.0. College Station, TX: Stata Corporation; 2007.

35. Rosenbaum PR, Rubin DB. The Central role of the propensity score in observational studies for causal effects. Biometrika. 1983; 70(1):41–55.

36. Raghunathan TE, Lepkowski JM, Van Hoewyk J, Solenberger P. A Multivariate Technique for Multiply Imputing Missing Values Using a Sequence of Regression Models. Surv Methodol. 2001; 27:1–20.

37. Greene, WH. Econometric Analysis. 2. New York: Macmillan; 1993.

38. Haddix, AC.; Teutsch, SM.; Corso, PS., editors. Prevention effectiveness: A Guide to decision analysis and economic evaluation. 2. Oxford: Oxford University Press; 2003.

39. US Food and Drug Administration. International Conference on Harmonisation; Guidance on Statistical Principles for Clinical Trials. 1998 Availability.

40. Drummond, MF.; Mcguire, A., editors. Economic Evaluation in Health Care: Merging Theory with Practice. Oxford: Oxford University Press; 2001.

41. Drummond, MF.; O'Brien, B.; Stoddart, GL.; Torrance, GW. Methods for the Economic Evaluation of Health Care Programmes. 2. New York, NY: Oxford Univeristy Press; 1997.

42. Gold, MR.; Russell, LB.; Siegel, JE.; Weinstein, MC., editors. Cost-effectiveness in health and medicine. New York: Oxford University Press; 1996.

43. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ. 1999; 18(3):341–364. [PubMed: 10537899]

44. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves–facts, fallacies and frequently asked questions. Health Econ. 2004; 13(5):405–415. [PubMed: 15127421]

45. Sendi PP, Briggs AH. Affordability and cost-effectiveness: decision-making on the cost-effectiveness plane. Health Econ. 2001; 10(7):675–680. [PubMed: 11747050]

46. Aos, S.; Lieb, R.; Mayfield, J.; Miller, M.; Pennucci, A. Benefits and costs of prevention and early intervention programs for youth. Seattle, WA: Washington State Public Policy Institute; 2004.

47. Bickman L. The functions of program theory. New Directions for Program Evaluation. 1987; (33)

48. Eisner M. No effects in independent prevention trials: can we reject the cynical view? J Exp Criminol. 2009; 5(2):163–183.

49. Sobel ME. Identification of Causal Parameters in Randomized Studies With Mediating Variables. J Educ Behav Stat. 2008; 33(2):230–251.

50. Raghunathan, TE.; Solenberger, PW.; Van Hoewyk, JV. IVEware: Imputation and variance estimation software user guide. Survey Methodology Program, Institute for Social Research, University of Michigan; Mar. 2002

Appendix A

Agency Record Review

For the economic analyses, an annual review of medical and other records provided detailed information on the services received across various service sectors. These reviews were conducted in each of 6 years from 1999 to 2005. (These years cover project years 9 through 14 for cohort 1, years 8 through 13 for cohort 2, and years 7 through 12 for cohort 3). The process of obtaining agency records can be summarized in four steps:

Step 1: Respondents Identify Agencies Providing Services

When parents or youth indicated receipt of services, they were asked for the name and address of the provider. Starting in the 1999 summer interview (year 9 for cohort 1 and year 7 for cohort 3) the respondent was asked to sign an authorization form to allow program staff to obtain agency records from any agencies they identified. Family authorization rates ranged from 99% of cases authorized in 1 year (at the Pennsylvania site) to 80% of cases authorized (at the Washington site). In addition to family authorization, successful agency identification relied on accurate information (agency name and location) provided from the families. (In instances where the agency could not be located because parents or youth had misidentified the provider, project staff would contact families asking for clarification.)

Step 2: Obtaining Cooperation of the Agencies Involved

If authorized, staff located the agency and invited them to participate in the study. Agencies were informed that the project would attempt to collect billing, medical or other records that might provide details on service type, service costs (amounts and payment sources), and number of days treated.

Of those cases for which we obtained family authorization and agency identification, agency participation rates ranged from 50% cooperation (first wave of data collection in Pennsylvania) to 90% cooperation (Durham, fourth wave of data collection). As much as possible, the project addressed agency concerns that might lead to refusal. For example, some agencies wanted the project to use agency-created authorization forms, and in those cases, we sought new signatures from the families. In some instances, agencies were still unable to participate because the records involved had been archived or were otherwise inaccessible. Overall, we were able to obtain records for roughly 60% of the services identified by parents or youth.

Step 3: Recording Services Information in a Database

Fast Track staff were trained to record agency information onto record review forms developed for the project. As one would expect, agency records varied widely in the amount of detail available. While some records would include very specific information broken down by service (e.g., distinct costs or diagnoses related to specific services), others would include overall information for the entire agency visit without any specificity at the service level. Record reviewers were trained to record agency data as it corresponded to 30 different service types (listed in Table A1). If less service-specific information was available, then data were coded using more general services types, like "medical office visit." Any services

recorded at highly detailed levels (e.g., cost for hospital sheets) were combined into more general service types (e.g., "inpatient services").

Another issue was timing. Agency data needed to be processed to correspond to the Fast Track project year since most other outcomes represent youth and family characteristics on an annual basis. A cut-off was selected to divide agency data to reflect this timing. That date was generally on or around July 1, so that a report of services from the agency data would be based on the annual "service delivery year" of July 1 through June 30. Services that were ongoing, especially those that involved inpatient status, would be divided into separate years based on that cut-off.

Particulars of service delivery also varied by agency. While we recorded service type information as it corresponded to the services listed in Table A1, other characteristics of the service delivery such as costs could not be distinguished at the service level, especially if presented in agency records as based on the overall admission rather than broken down at the level of service. In such cases, costs and/or diagnoses were recorded as they corresponded to the first/primary service listed (e.g., individual counseling) with indication in the data for the other services that costs were recorded under that primary service.

Step 4: Data Processing for Economic Evaluation

For the economic evaluation, we wanted to estimate (i) the total dollars spent on services across years, and (ii) the number of certain service types (e.g., number of mental health services) received across all study years. This involved generating totals on these variables across all services, agencies, and years for each subject. In order to derive the count of certain service types, we created categorical indicators for each service type coded from the agency records (indicated in the footnote to Table A1) based on nature of the service: mental health, general health, medication, drug-alcohol treatment, juvenile justice service, and pregnancy-related. (In some instances, service type was dictated by the agency category.)

Missing Data

Before creating summary variables from the full data, we needed to address missing data characteristics of the medical records. As noted, agency data may be missing for any of several reasons. First, the family could refuse to authorize record review. Second, an agency might be unwilling or unable to provide the needed information. Even if an agency was accessed and cooperative, subject records may still contain missing fields because of incomplete records. As noted above, roughly 40% of SACA-reported agencies were missing record review information due to various complications in the data collection process.

Given the unique nature of the agency data, we carried out data imputation for medical records separately from the other outcomes. Moreover, it was important to impute separate data sets by service sector in order to most effectively model the missing data process using the most relevant information for the services outcomes (within similar agencies and for similar needs). Data sets for these models included the subjects' annual cost totals, number of service types, and number of days served at each family-reported agency for each Fast Track project year. This would allow the agency medical records to be combined with the family-report information (SACA's) that had been provided for each agency for any year services were reported. The information from the families was used to help set up imputation models for the missing agency medical records since family information on an agency visit would be non-missing in cases where we were not able to complete the agency record review. The following figure demonstrates the structure of the data when family information was provided but no successful record review occurred. In such cases,

imputation models relied on data provided by the families (left columns in the example table below).

Table A1

Fast Track Record Review form Service Types

Residential:
Inpatient services [a,b]
Residential treatment [a]
Detention/jail [e]
Group home services [a]
Shelter services
Intensive Outpatient:
Alcohol & drug OP [d]
In-home services/mobile therapy [a]
Day treatment/partial hospitilization [a]
Other Outpatient:
Individual counseling [a]
Group counseling [a]
Family counseling [a]
Case management
Respite services
Collateral therapy [a]
In-school academic counseling
In-school emotional/behavioral counseling [a]
Psych.testing/assessment [a]
A&D testing (including urine screens) [d]
Medical:
Medications mgmt [c]
Lab [b]
Pharmacy [c]
Medical office visit [b]
ER services [b]
Pre- or post-natal services/obstetrics [f]
Surgery [b]
Other Services:
Foster care
Travel/transportation
Other
Misc. cost codes:
Capitation/case-rate
No show (didn't receive service but was billed)

Note: Services were categorized into the following general service types:

[a] mental health;

[b] general health;

$c$medication;

$d$drug-alcohol treatment;

$e$juvenile justice service;

$f$pregnancy-related

Agency medical records were imputed using IVEware (Raghunathan[50]). As indicated, we divided the medical records database into separate data sets based on the type of service sector so that the nature of service delivery was similar across cases. This distinction was carried out using the service sector *as indicated by the respondent* in the SACA since those data had complete information on the service sector of the agency visit (i.e., the SACA information was non-missing by design since it had to exist in order to instigate agency follow-up). The following designates the partitions of the data based on the respondent characterization of the agency (i.e., where in the SACA the agency was reported). Separate data sets were imputed based on these distinctions:

- Juvenile justice agency/detention center

- General health (General hospital or emergency department)

- Social service (including emergency shelter, foster home or respite care)

- Outpatient mental health (including day treatment center, substance abuse clinic, in-home provider, mental health center)

- Pediatrician/family doctor

- Inpatient mental health facility (psychiatric hospital, group home, residential treatment center)

- Outpatient therapist

Table A2

Example of Services/agency Information Available for a Particular Project Year

	Family information (SACA)			Agency information		
SubjectID	Service sector category	Agency name	# days admitted across year (family report)	Service type	Service costs	# days
0001	Outpatient therapist	Jones and Assoc. Family counseling	10	–missing–	–missing–	-missing-

Data were also separated by intervention status before imputations. Subjects from the non-high-risk/normative sub-sample were randomly split between intervention groups in order to provide more information for the missing data model estimation. Imputation bounds (variable ranges) were set based on frequency distributions for non-missing cases. We imputed five data sets for each service sector data file. The following variables were included in the imputation process; the services variables were based on the full year of service (for each agency reported):

- The 25 baseline covariates

- Service type counts for agency-year from the record review (e.g., number of mental health services)

- Service delivery characteristics from the record review (bill amount per day, number of days served, project year of service)

- Other study background characteristics (Study site, risk-status, cohort, whether African-American, gender)

- Information provided on the service from the family (number of days admitted, agency category)

Skewed variables from agency data collection – namely billing amount and number of days served – were square-root transformed before imputations.

Separate service sector data sets were combined after imputations and agency cost totals were calculated post-imputation by multiplying the bill-per-day rate by the number of days served at that facility. Agency amounts were then summed across years for each subject and merged with the full Fast Track sample. Service outcomes were set to zero for any subject who had no report of service receipt (from parent or child-SACA's) during the measurement period but did participate in the administrations of the SACA. This assumes that non-reporting of services by families is valid information and that, indeed, the youth received no services and spent zero dollars on agency costs across the 6 years. Families who dropped out of the Fast Track study before the SACA was administered were not included in statistical analyses of the medical records given the difficulty in imputing what agency a person would be admitted to, which is necessary information in order to impute record review outcomes. The final services data included total services information across years per-subject, combining both agency data and respondent-report of service delivery.

Appendix B

Table B1

Year-1 (baseline) measures

Variable(s)	Instrument
Emotion Recognition Questionnaire: Were children able to correctly identify feelings?	Emotion Recognition Questionnaire
Feeling Scale – Depression: List of frequency and severity of symptoms of depression mother has experienced in last week	Feeling scale
Percentage of hostile attributions child makes when asked why another child behaved as he or she did	Home Interview With Child hostile attributions
Number or percentage of retaliatory and / or punitive responses a child generates to solve social problems	Home Interview With Child punitive retaliations
Appropriateness Score: Percentage of appropriate solutions child generates in response to social problems	Interview of Emotional Experiences
IPE - Family Satisfaction: How satisfied is mom with the social and instrumental support she receives from family members	Inventory of Parent Experiences:
IPE - Friendship Satisfaction: How satisfied is mom with the social support she receives from friends	Inventory of Parent Experiences
How warm was the mother toward the child during an observed interaction task	Parent Child Interaction Task: warmth
Physical punishment: Number of vignettes of typical child problems in which mother advocated use of physical punishment	Life Changes
Stress scale: Number of stressful life events family encountered last year	Life Changes
Number of vignettes of typical child problems in which mother said she would yell at child or verbally reprimand him or her	Life Changes
Socioeconomic Status Continuous Code: Combination of parent educational achievement and occupational prestige	Family Information Form

Variable(s)	Instrument
Tally over three separate days of whether different oppositional or aggressive behaviors occurred	Parent Daily Report
Self-rating of parenting practices	Parent Questionnaire: discipline construct
Child-friendly Interior Interviewer rating of presence of books, toys, room to play, and so forth	Post-Visit Inventory
Home Environment Interviewer rating of how safe and clean home is	
Parent rating of child's prosocial skills and emotion regulation skills	Parent Social Competence
Percentage of competent responses child can generate to solve social problems	Social Problem Solving
Teacher rating of oppositional and aggressive behavior; used as first-stage in screening process	Teacher Observation of Child Adjustment
Externalizing T Score Combination of delinquent and aggressive behavior syndromes on the Teacher's Report Form, which is teacher version of Child Behavior Checklist	Teacher report form
Hyperactivity scale of the Teacher's Report Form	
Inattention scale of the Teacher's Report Form	
Word identification score of Woodcock-Johnson test of academic achievement	Woodcock-Johnson
Parent rating of the quality of the neighborhood	Parents were asked some questions about neigh
Short form of child IQ	WISC = IQ test

Table 1

Fast Track Measures for Economic Evaluation

Domain	Project years	Source
Inpatient services use (amount, whether)	7 through 13	Services Assessment for Child and Adolescents-modified (SACA): Parent version
Outpatient services use (amount, whether)		
General health services use (amount, whether)		
General health services use for behavioral health purposes (amount, whether)		
Arrest history (amount, whether)	9 through 13	
Court appearances (amount, whether)		
Police contact (amount, whether)		
Detention center stays (amount, whether)		
Inpatient services use (whether)	11 through 13	Services Assessment for Child and Adolescents-modified (SACA): Youth version (Youth)
Outpatient services use (whether)		
Arrest history (amount, whether)		
Court appearances (amount, whether)		
Police contact (amount, whether)		
Detention center stays (amount, whether)		
Found guilty of a crime (whether)		
Whether medication for behavior/attention	2 through 13	Life Changes inventory
Antisocial behavior subscale (8yr–13yr)	8 through 13	Self-Report of Delinquency (Youth)
Severe crimes committed (amount, whether)	Year 13 (cumulative)	Court records inventory (Records)
Any crime committed (amount, whether)		
Age at first crime committed (hazard model)		
Jail stays (whether)		
Repeated a grade (whether)	2 through 13 (except graduation, year 13 only)	School records inventory (Records)
Special education use (whether)		
Graduation from high school (whether)		
Time until repeating a grade (hazard)		
Time until requiring special education services (hazard)		
Smoking in past month	8 through 13	Tobacco, Alcohol, and Drugs (Youth)
Drunk on alcohol in past year		
Marijuana use in past month		
Substance abuse subscale	9 through 13	Parent Daily Report
Public assistance dollars received (amount, whether)	7 through 12	Family Information Form
Public assistance dollars received (amount, whether)	13	Income and Employment (Youth)

Note. Measurement source is parent/primary caregiver unless otherwise noted in the Source column.

Table 2

Results of Regression Analyses for Fast Track Project Economic Outcomes

Measure	Source	Model type	Overall impact (pooled estimate)	Interpretation	Significant subgroup intervention effects (and empirical Bayes point estimate)
Health and Mental Health Services					
Whether used inpatient services	Parent	L	0.08	N/A	None
Whether used outpatient services	Parent	L	−0.077	N/A	None
Whether received mental health services from a general health provider	Parent	L	−0.344**	Reduced likelihood of MH services from a GH provider	PML (−0.51) PFH (−0.56)
Whether used general health service	Parent	L	−0.171*	Reduced likelihood of general health services	NFL (−0.27)
Number of inpatient services	Parent	N	0.316#	Increased use of inpatient services	NML (.90) NFH (.94) PFH (.65)
Number of outpatient services	Parent	N	−0.122	N/A	NML (−0.39) NMH (−0.40) PML (−0.31)
Whether used inpatient mental health services (years 10–13)	Parent	L	0.046	N/A	None
Number of inpatient mental health services (years 10–13)	Parent	N	0.116	N/A	NML (0.87)
Whether used outpatient mental health services (years 10–13)	Parent	L	−0.250**	Reduced likelihood of outpatient MH services	WMH(−0.88)
Number of outpatient mental health services (years 10–13)	Parent	N	−0.210**	Reduced number of outpatient MH services	PMH (−0.41)
Number of mental health services through general health providers	Parent	N	−0.336	N/A	NML (−1.07) PML (−0.98) WML(−1.28) PFH(−1.48)
Number of general health services	Parent	N	−0.160**	Reduced use of general health services	DMH(−0.74) WFL(−0.26) NML(−0.29) PFH (−0.33) NMH (−0.26)
Whether outpatient MH services (in past 2 years)	Youth	L	−0.357**	Reduced use of outpatient MH services	None
Whether inpatient MH services (in past 2 years)	Youth	L	−0.272	N/A	None
Whether medication for behavior/attn	Parent	L	0.127	N/A	DFL(0.97) DML(1.19) PFH (−0.67) DMH(0.58)
Delinquency and involvement in Juvenile Justice System					
Number of severe crimes (through year 13)^	Court Records	N	0.103	N/A	None

Measure	Source	Model type	Overall impact (pooled estimate)	Interpretation	Significant subgroup intervention effects (and empirical Bayes point estimate)
Number of crimes including less severe offenses (through year 13)^	Court Records	N	−0.113	N/A	None
Whether severe crime (through year 13)^	Court Records	L	−0.004	N/A	None
Whether any crime including less severe offenses (through year 13)^	Court Records	L	−0.234	N/A	None
Whether ever in jail^	Court Records	L	−0.460#	Reduced likelihood of being in jail	None
Age at first crime committed (severe crime)	Court Records	S	−0.092	N/A	None
Age at first crime committed (any crime)	Court Records	S	−0.122	N/A	None
Number of arrests in past year^	Parent	N	0.019	N/A	DH(−0.70)
Number of court appearances in past year^	Parent	N	0.12	N/A	NL(0.48) DH(−0.58) WH(0.41)
Number of police contacts in past year	Parent	N	0.194	N/A	PFL(1.0) PFH(0.85) DMH(−0.59)
Number of nights in detention center in past year^	Parent	N	0.089	N/A	DH(−0.40) PH(−0.47)
Whether arrested in past year^	Parent	L	0.057	N/A	None
Whether court appearance in past year^	Parent	L	0.153	N/A	None
Whether police contacts in past year	Parent	L	−0.051	N/A	None
Whether spent time in detention center^	Parent	L	−0.012	N/A	None
Were you found guilty or convicted of any crime?^	Youth	L	−0.11	N/A	None
Number of arrests in past year^	Youth	N	−0.318**	Decreased number of arrests	PH(−0.45)
Number of court appearances in past year^	Youth	N	−0.211	N/A	NL(−0.38) PH(−0.52)
Number of police contacts in past year	Youth	N	0.178	N/A	PML(0.84)
Number of nights in detention center in past year^	Youth	N	−0.002	N/A	PH(−0.47)
Whether arrested in past year^	Youth	L	−0.103	N/A	None
Whether count appearance in past year^	Youth	L	−0.134	N/A	None

Measure	Source	Model type	Overall impact (pooled estimate)	Interpretation	Significant subgroup intervention effects (and empirical Bayes point estimate)
Whether police contacts in past year	Youth	L	0.049	N/A	None
Whether spent time in detention center^	Youth	L	−0.029	N/A	None
Count of antisocial behaviors	Youth	N	−0.067	N/A	DFL(−0.50) NFL(0.32) DML(0.27) PML (−0.52) WML(0.46) PMH(−0.51)
Whether antisocial behavior	Youth	L	−0.181 **	Decreased likelihood of anti-social behavior	PML(−0.46)
School Failure/School services					
Whether repeated a grade	School Records	L	0.112	N/A	None
Whether special education services	School Records	L	0.107	N/A	DFL(−0.96) PML(1.06) PFH(0.78) WFH(0.48)
Whether graduated from high school	School Records	L	0.211	N/A	None
Time until repeating a grade	School Records	S	0.015	N/A	None
Time until requiring special education services	School Records	S	0.129	N/A	None
Substance Abuse					
Number of days smoked in past month	Youth	N	−0.045	N/A	PFH(0.63)
Number of days very drunk in past month	Youth	N	0.117	N/A	None
Number of times use marijuana in past month	Youth	N	0	N/A	DML(0.53)
Substance Abuse (ParentDailyReport)	Parent	N	−0.112	N/A	PML(0.34) DMH(−0.47) WMH(−0.39)
Financial Assistance					
Amount of public assistance dollars received (years 7–12)	Parent	N	2.28	N/A	None
Whether public assistance dollars received (years 7–12)	Parent	N	−0.013	N/A	None
Amount of public assistance dollars received (year 13)	Youth	N	0.54	N/A	None
Whether public assistance dollars received (year 13)	Youth	N	0.063	N/A	None

Note: N = Negative binomial regression; L = Logistic regression; S = Survival models

Initials for sub-group: D = Durham; N = Nashville; P = Penn; W = Wash; F = Female; M = Male; H = Higher Risk; L = Lower Risk

Statistical significance:

#
p<.10 /

*
p<.05 /

**
p<.01

^Tests done on eight study sub-groups

Table 3

Marginal Effects of Significant Group Differences

Intervention-Control group difference (if p<.05)	Source	Model type*	Marginal effect	Interpretation**
Whether received mental health services from a general health provider	Parent	L	−0.080	Intervention youth 8% less likely to receive mental health services from general health provider than the control group
Whether used general health service	Parent	L	−0.161	Intervention youth 16% less likely to receive general health services than the control group
Number of general health services	Parent	N	−1.136	Intervention youth have 1.14 fewer visits to a general health provider than the control group
Whether used outpatient mental health services (years 10–13)	Parent	L	−0.143	Intervention youth 14% less likely to receive outpatient mental health services
Number of outpatient mental health services (years 10–13)	Parent	N	−0.641	Intervention youth have .64 fewer visits to a outpatient mental health provider than the control group
Whether outpatient MH services (in past 2 years)	Youth	L	−0.236	Intervention youth 24% less likely to receive outpatient mental health services
Number of arrests in past year	Youth	N	−0.755	Intervention youth have .76 fewer arrests than the control group
Whether antisocial behavior	Youth	L	−0.200	Intervention youth 20% less likely to exhibit antisocial behavior than the control group

*
L = Logistic; N = Negative binomial

**
Based on a given year (on average across years where outcome measured)