Published in final edited form as:

Child Maltreat. 2011 May ; 16(2): 114–125. doi:10.1177/1077559511402048.

Change Trajectories During Home-Based Services With Chronic Child Welfare Cases

Mark Chaffin¹, David Bard¹, Debra Hecht¹, and Jane Silovsky¹

¹University of Oklahoma Health Sciences Center, Department of Pediatrics

Abstract

This study examines how risk factor change patterns vary with case chronicity, and whether risk factor improvement still predicts lower recidivism risk among chronic cases. 2,175 parents in home based child welfare services were surveyed for risk factors at pre-treatment, post-treatment and 6-month follow-up. Mixture modeling of latent difference scores identified change trajectory classes related retrospectively to chronicity and prospectively to recidivism. Five change trajectories were identified: stable low problem, stable high problem, sustained improvement, relapsing, and paradoxical. Chronicity was associated with a decreasing probability of membership in the stable low problem trajectory and increasing probability of membership in the stable high problem and sustained improvement trajectories. Cases with more favorable trajectories recidivated less across levels of chronicity. Findings suggest that chronic cases may improve little, but still retain a stable or increasing chance of sustained improvement associated with lower risk. A cumulative service benefit might be one possible explanation for this observation, and might suggest that repeated intervention efforts are not always wasted on chronic cases. The current episodic and reactive service delivery model in child welfare may be a mismatch with chronic cases where progress is absent or tends to occur cumulatively across service episodes.

Keywords

child welfare; chronic cases; recidivism; home based services

Child welfare is increasingly becoming a system that serves chronic cases, particularly chronic child neglect cases. Over the past two decades, incidence rates of cases entering child welfare for neglect have held relatively constant, while rates of physical and sexual abuse cases entering child welfare have declined (Finkelhor, Jones, & Shattuck, 2010). As we might expect from these incidence patterns, the proportion of all child welfare cases that are neglect cases is steadily increasing and neglect currently comprises 71% of cases nationally (U.S. Department of Health and Human Services/Administration on Children,

Reprints and permission: sagepub.com/journalsPermissions.nav

Corresponding Author: Mark Chaffin, University of Oklahoma Health Sciences Center, P.O. Box 26901, Oklahoma City, OK, USA, Mark-chaffin@ouhsc.edu.

The opinions expressed are those of the authors and do not necessarily reflect those of the NIMH or the CDC.

Declaration of Conflicting Interests

The author(s) declared no conflicts of interest with respect to the authorship and/or publication of this article.

Youth and Families, Children's Bureau [USDHHS/ACYF], 2010). Neglect is a more recurrent form of maltreatment, and therefore child welfare caseloads are progressively becoming comprised of chronic neglect cases, including cases that may have received the same or similar services multiple times in the past. For example, one state study of child welfare home-based services reported that among families entering services, the count of previous child welfare reports grew from a median of one in 2001 to a median of three in 2009; and the percentage of entering cases with chronic child welfare histories (defined as four or more previous reports) almost doubled over the same time frame from slightly under 20% to almost 40% in 2009 (Hecht, Chaffin, & Silovsky, 2010). Extrapolating these trends, it is possible that chronic neglect cases may soon characterize half or more of the total child welfare service population.

Chronic child welfare cases have a high recidivism rate (DePanfilis & Zuravin, 1998; Fluke, Yuan, & Edwards, 1999). Accumulated prior report count is one of the stronger, if not the single strongest, predictor of future recidivism and there also can be temporal acceleration of inter-event intervals associated with chronicity, with each subsequent new report occurring on average faster than the prior report (English, Marshall, & Orme, 1999). Chronicity is relevant to downstream child outcomes (Widom, 2009). Several studies point to chronic maltreatment as being consistently associated with poorer child outcomes (English et al., 2005; Ethier, Lemelin, & Lacharite, 2004; Lemmon, 2006) and the form of these poorer child outcomes can vary depending on which developmental period and how many different periods overlap with maltreatment (Graham et al., 2010). Chronic maltreatment may impact children directly, or may be a proxy for multiple comorbid family problems, given that families that are chronically in the child welfare system often experience problems across socioeconomic, social, interpersonal, and mental health domains (Nelson, Saunders, & Landsman, 1993).

Families chronically entering child welfare have had opportunity to experience multiple service episodes, potentially including receipt of the same or very similar services multiple times in the past. A key question is whether continuing to offer services to chronic cases is a wise strategy and whether chronic families retain any reasonable prognosis for positive change and reduced recidivism over the course of these services. Little is known about patterns of risk factor or problem change during services among chronic cases and whether different change patterns might predict lower or higher risk for recidivism. The purpose of this study was to describe patterns of change during and after home-based child welfare services and examine how change patterns are associated retrospectively with case chronicity and prospectively with risk for future recidivism.

There are a number of nonexclusive hypotheses that might explain how change during services, chronicity, and recidivism risk might be related. Each hypothetical also might have a different set of implications for service system structure. One hypothesis is that some families who are chronically in child welfare have substantial and intractable problems that respond little to services. This pattern might correspond to what David P. H. Jones described as "the untreatable family" in his classic article of the same name (Jones, 1987). We might predict that this pattern would be reflected in a change trajectory characterized by high initial problems and insufficient change across the service interval and beyond and that

membership in this change trajectory class would be more prevalent among more chronic cases and would be prospectively associated with higher recidivism. Jones (1987) discussed the implications of untreatability on child safety, staff burnout, and depletion of system resources. One implication is that there might be little point in triaging rehabilitative effort to untreatable families and that service resources would better be allocated elsewhere. Alternately, this pattern might suggest the need to radically reconsider service goals, such as refocusing on harm reduction (e.g., alternative care-givers, enriched day care, respite, and co-parenting) rather than wasted rehabilitation efforts.

A second hypothesis is that families who are chronically in child welfare may improve during services but lose or fail to sustain that improvement over time. This is a relapsing pattern. A relapsing pattern has implications that are very different from those associated with untreatable families. This pattern would suggest greater attention devoted to how improvements can be consolidated or maintained over time, either by procedures such as booster sessions or follow-up care. With any chronic condition (e.g., diabetes, asthma, drug dependency, and obesity), long-term maintenance issues are at least as critical to success as achieving initial improvement, a perspective that has been particularly influential in rethinking contemporary services for drug dependence (McLellan, Lewis, O'Brien, Hoffman, & Kleber, 2000). The relapse hypothesis would predict that a high problem/good improvement/relapse trajectory would emerge from the data and that membership in this trajectory would be more prevalent among chronic cases and prospectively associated with higher recidivism.

A third hypothesis is that families get worse over the course of child welfare services. This paradoxical pattern could result in a downward spiral or vicious cycle of services and recidivism. It is not unheard of that services with plausible logic models, good intentions, and presumptive benefits may ultimately prove to be harmful (Lilienfeld, 2007) or impose more stresses than they alleviate. Inferring deterioration over the course of services is challenging because measured deterioration may reflect increases in willingness to report problems rather than actual deterioration. The paradoxical pattern scenario would predict that a trajectory would emerge from the data involving increasing problems from baseline to posttreatment, after controlling for changes in willingness to report, which is then followed by partial improvement or stabilization after services are withdrawn. We would expect that this trajectory would be increasingly prevalent among chronic cases and prospectively associated with higher recidivism.

A fourth hypothesis is that chronicity can be observed under cumulative patterns of change. One version of the cumulative hypothesis is probabilistic—that is, that high-problem families have a modest but stable probability of improving during any given service episode. Because the probability of showing sustained improvement would be modest, only a few high problem cases would improve and stay improved during any given episode. Cases might require multiple episodes before services would eventually “hit” in a probabilistic sense. Another version of the cumulative change hypothesis is dose related. Some cases may simply require multiple service encounters before accumulating a sufficient dose. For example, some parents might need to practice change skills multiple times before finally achieving full stable change. This sort of pattern has been observed in the addictions

literature (Hser, Grella, Chou, & Anglin, 1998), where it is now accepted that individuals should be encouraged to “try treatment again” in the face of incomplete recovery or even multiple relapses. Under this scenario, chronicity does not necessarily imply long-term untreatability and does not vacate the wisdom of offering the same services over and over again. Under this scenario, we might predict observing a sustained positive change trajectory (high problems that decrease during services and remain low at follow-up). This trajectory would have only a modest probability of occurring, but its probability would remain stable across levels of chronicity (e.g., a cumulative odds pattern) or possibly even increase (e.g., a cumulative dose pattern). The cumulative benefit hypothesis also would predict that when a sustained improvement trajectory is observed, recidivism risk should decrease and decrease fairly uniformly across levels of chronicity—that is, chronicity would not moderate the relationship between trajectory class membership and lower recidivism.

Other change trajectories might be predicted to be negatively associated with chronicity based on previous findings. For example, because chronic cases often are high problem cases (Kaplan, Schene, DePanfilis, & Gilmore, 2009; Nelson et al., 1993), we would expect to observe fewer cases with initially low problem levels as the level of chronicity increases.

The strategy this study used to inform these various possibilities can be described in three conceptual steps. First, latent class or mixture modeling of multivariate risk and problem variables was used to test whether a single parametric trajectory or unobserved multiple trajectories best fit the data, and if distinct trajectories fit the data well, to classify parents according to which change trajectory best fit their observed data. Second, trajectory membership was retrospectively tested for association with case chronicity (i.e., accumulated prior system entries) across the previous decade, modeling how the probability of trajectory class membership changed with increasing chronicity. Third, trajectory membership was prospectively tested for predicting child welfare recidivism hazard across an approximately three year average follow-up, including prediction of recurrent (i.e., more chronic) recidivism events. We then compared the overall pattern of observations to the patterns predicted by the hypotheses described above.

Four constructs were examined for change. These were selected for their risk relationship with maltreatment, their amenability to change over time, and to represent somewhat distinct risk domains—mental health, basic concrete resources, social support, and actuarially based parenting risk. The first problem construct was parental depression. In a representative population sample, parental depression was identified as one of the stronger prospective risk factors for the initial onset of child maltreatment. Compared to nondepressed parents, depressed parents were around three times more likely to begin maltreating their children (Chaffin, Kelleher, & Hollenberg, 1996). A history of prior mental health problems, including depression, is also associated with child welfare recidivism (Drake, Jonson-Reid, & Sapokaite, 2006). The second construct was concrete resources. Perhaps, the dominant demographic characteristic of families in child welfare is poverty and families with lower socioeconomic and family resources are many times more likely to enter child welfare (Drake & Zuravin, 1998). Basic resource challenges are a particular problem among chronic neglect cases. The third construct was social support. Low social support is common among parents in child welfare and is a factor in maltreatment recurrence risk (DePanfilis &

Zuravin, 1999). The fourth change measure was the Child Abuse Potential Inventory, which is a tool developed specifically to measure abuse risk and focuses mostly on parenting attitudes, distress, personal characteristics, and family characteristics that actuarially predict risk.

Method

Participants

Participants in the study were 2,175 parents who were enrolled in a statewide network of home-based contracted family preservation and support programs operated by large nonprofit community-based agencies, one for each of the six child welfare administrative regions of the state. All participants were parents or caregivers referred by child welfare to the programs due to reports that they committed physical abuse and/or neglect of children in their household. Parents receiving services due to child sexual abuse were excluded from the study because these cases were felt to present distinct services issues and needs. Parents were recruited for the research in their homes by a research assistant shortly after service enrollment and were provided with a \$25 gift certificate at each data collection wave as compensation. Recruitment and informed consent procedures were approved by the University institutional review board (IRB), and participant welfare was overseen by a Data and Safety Monitoring Board that included representatives from child welfare, the provider agencies, independent health professionals and researchers, and an independent expert on research with culturally diverse samples. Study participation involved data collection only and did not alter the services received by the family or the child welfare case disposition. Only one parent per household was enrolled, with first priority given to the parent identified as the primary caregiver. About 3,116 prospective participants were approached, 18 did not complete the recruitment process, 23 were determined to be ineligible, and 816 declined to enroll or complete baseline data collection, yielding an overall enrollment of 2,259 (72% of all individuals approached). About 84 participants were withdrawn after enrollment (50 voluntarily and 34 involuntarily), yielding the analyzable sample of 2,175. Individual data on non-enrollees was unavailable. However, information on over 5,000 participants in these same programs from adjoining time periods was available and their demographic characteristics were comparable to the study sample (88% female; 67% non-Hispanic White, 13% American Indian, 12% African American, 5% Hispanic, and 2% Other; median age = 30).

In the study sample, 91 percent of the 2,175 participants were female with a mean age of 29 years ($SD = 8$; range = 18–75). Participants had a median of three children in their family, 76% had at least one preschool age child, and 8% of women reported being pregnant at baseline. About 27% lived in an urban setting, 63% lived in small communities, and 10 % lived rurally. Residential instability was common, with 52% having lived in their current community less than 3 years and 54% having moved more than twice in the last 5 years. About 67% of participants were non-Hispanic Whites, 16% were Native American, 9% were African American, 5% were Hispanic, 0.4% were Asian, and 2.3% indicated another race/ethnicity or did not answer. About 31% were married, 15% were cohabitating, 14% were separated, 16% were divorced, 2% were widowed, and 23% were never married. About 40%

had less than a high school education, 39% had a high school diploma or equivalent, 16% completed some college, and 4% had completed college. Median household income was \$930/month. Applying current U.S. federal poverty line criteria for income and family size, 83% of households fell below the federal poverty line at baseline. About 27% indicated that they were currently unemployed, 26% were homemakers, 29% had a full-time job, 6% were students, and the remainder indicated part-time or self-employment. The mean baseline score on the Beck Depression Inventory (BDI) was 13 ($SD = 12$, range = 0–61), and applying a cutoff score of 19, 25% would be considered to have a clinically significant symptom level. The mean baseline score on the Child Abuse Potential Inventory was 162 ($SD = 106$, range = 1–442), which is approximately equal to the Child Abuse Prevention Inventory signal detection cutoff score for detecting physical abusers, and is well above normative mean of 91 (Milner, 1986).

Participants had a mean of three and a median of two unduplicated prior child welfare referrals (range of 0–30; $SD = 2.7$). About 87% of all prior referrals were for child neglect. As child welfare chronicity increased, the chances of having at least one prior neglect report became a virtual certainty—96.5% of all cases with two prior referrals had at least one neglect report, as did 99.9% of cases with three or more past referrals. About 30% had at least one child placed outside their home at baseline, and for these cases, services were normally part of a reunification plan.

Procedures

Data for the study were collected during 2003–2006 in participants' homes by independent research assistants using Audio Computer Assisted Self-Interview (ACASI). An initial set of ACASI practice items was included at the beginning of the interview for the research assistant to demonstrate the system and establish that parents understood the system and the test items. Parents had the option to complete the interview with or without audio. If chosen, the audio option read each question and each response option as it was highlighted on screen. Parents gave responses by touch screen. Computer interviews were conducted while the research assistant waited or supervised the children in order to provide the parent with uninterrupted private time to respond to items. Only research assistants and not home visitors were involved in data collection. Research assistants normally did not view parents' responses unless the parent requested assistance. A federal Certificate of Confidentiality was obtained and no individual research data was shared with child welfare authorities or service providers. Measures were collected at baseline (i.e., around service entry), around the end of the services (median time = 205 days from baseline; $SD = 92$; $n = 1,279$), and again at around 6 months after service exit for post-program follow-up (median = 405 days from baseline; $SD = 87$; $n = 892$). The dominant reason for interwave attrition was participants who could no longer be located, despite multiple attempts to follow up using both official and unofficial contact sources (e.g., participants who moved and left no forwarding address, who had become incarcerated, or for whom no current location could be obtained from the contact persons they identified, the home visiting service agency, or child welfare).

Measures

Beck Depression Inventory-2 (BDI-2)—The BDI-2 (Beck, Steer, & Brown, 1996) is a 21-item multiple-choice self-report questionnaire designed to measure symptoms of depression. Published internal consistency of the scale is .93, and test–retest stability is .93 (Beck et al., 1996). Observed alpha in the study sample was .94.

Family Resources Scale–Revised—The FRS (Dunst & Leet, 1987) is a 40-item self-report scale designed to measure the adequacy of basic concrete needs in households with children. FRS items are ordered as a hierarchy of basic needs drawn from an ecological perspective and including very basic needs (e.g., having enough food, having shelter, and clothes); social needs (e.g., having enough time with family and having time for friends); needs involving transportation, medical and dental care; and finally less critical needs such as having sufficient resources for extras, entertainment, savings, and so on. Examining the raw data, response patterns did not follow a clear hierarchical pattern or show a clear ceiling, so simple summative rather than ceiling scores were used. The summative score reflected the mean degree to which the full range of needs was met. Observed alpha for the overall scale in the study sample was .85, supporting summative scoring.

Social Provisions Scale (SPS)—The SPS is a measure of perceived social support (Cutrona & Russell, 1987). SPS items are drawn from six aspects of social support (friends, family, etc.). SPS internal consistency estimates range from .83 to .94 (Mancini & Blieszner, 1992). Observed alpha in the study sample was .84.

Child Abuse Potential Inventory (CAPI)—CAPI (Milner, 1986) is a 160-item agree/disagree format parent self-report questionnaire developed to estimate abuse risk. Item content is related primarily to parenting stress and attitudes, relationship conflict, emotional distress, and parent characteristics. The CAPI Abuse Scale, which is the main scale of the CAPI, has been reported to have high internal consistency ($KR-20 = .92$ to $.95$), a 1-month test–retest stability of 0.83 (Milner, 1986), and actuarial predictive validity for discriminating maltreating from non-maltreating parents and predicting future child welfare reports (Chaffin & Valle, 2000; Milner 1986). Alpha for the Abuse Scale in the current study sample was .92. The CAPI also includes an 18-item Lie Scale measuring social desirability response bias. Items on the Lie Scale reflect denial of minor but socially undesirable faults to which people will ordinarily readily admit. The scale has been found to correlate significantly with other general social desirability measures and to discriminate between parents instructed to answer honestly versus to answer in a socially desirable manner (Milner, 1986). The observed alpha for the CAPI Lie Scale in the study sample was .78. Pretesting construct validity of the Lie Scale revealed significant negative correlations with symptom and risk measures (BDI and CAPI), and significant positive correlations with strength measures (FRS and SPS), as expected. CAPI Lie Scale scores diminished over time, suggesting changes in willingness to report (i.e., time dependent response bias), and therefore all subsequent models corrected observed scores for response bias using the corresponding CAPI Lie Scale score. This was felt to be important so that change trajectories would better reflect construct change, not change in response bias.

Prior child welfare reports and recidivism—The state child welfare agency maintains a central database for child welfare data. The database had been in place for a decade prior to the start of the study and retained records of all reports. Matches were executed to identify all prior reports involving the study subject as the perpetrator using a combination of database identifiers at the family and individual level, additionally verified by date of birth match. Reports and allegations that were ruled out or screened out were excluded. Reports involving an alternative response (or differential response) disposition rather than investigation disposition were retained. The extracted events and allegations within events were aggregated across maltreatment types, across children in the family and across report dates in order to yield unduplicated counts and temporal sequences. Mean follow-up time for future reports was 1,048 days (range = 431–1,554 days). Future report survival times were adjusted by removing risk deprived time intervals. Risk deprived intervals are those where official child placement data sources indicated that no children were present in the home (i.e., all children were in out-of-home placement), and therefore the parent was not actively at risk for a recidivism report. Pregnant participants were not treated as risk deprived. Surveillance reports (i.e., reports made by the home visitor) accounted for 5% of all recidivism incidents and were retained in the recidivism data, given that they were infrequent, all participants were subject to a similar level of service related surveillance, and no group comparisons that might be biased by differential surveillance were planned.

Services—Family preservation and support services were delivered in the home by bachelor's level home visitors ($n = 229$), supervised by licensed master's level clinicians, and employed by community-based agencies, one for each of the six administrative regions in the state, under contract with child welfare. The expected home visiting service dose was 6 months, and this network of home visiting programs was designed to offer the most intensive services currently available within the state child welfare system. The planned frequency of visits varied, but was required to be at least weekly, and was designed to be more frequent during early weeks of the service episode. Services were designed to prevent foster care placement or to promote and stabilize reunification from foster care. Service content was specified by state contract and included several required elements. Required basic service elements included case management and linkage to outside services, direct assistance with parenting problems, assistance with meeting basic needs, direct assistance solving family problems and conflicts, basic family violence safety planning, monitoring children's welfare in the home, crisis management, and providing support. Participants were screened for mental health, domestic violence, and social services needs using standardized tools and linked to community resources. Each home visitor had access to a \$500 budget to assist families in meeting basic concrete child care related needs (e.g., getting utilities turned on in the home). Because most cases involved child neglect, services focused mainly on helping parents to create a physically adequate home environment, promoting household stability, and improving basic caregiving and parenting. Services differed across agencies in some aspects, including curricula used to address goals and quality control techniques, but all programs shared a comparable contractual, management, staffing and referral framework and used a common set of reporting and screening tools.

General statistical modeling considerations—All structural models were estimated using Mplus 6.0 software (Muthén & Muthén, 2010). Scores on all psychometric measures were transformed by binning and/or simple mathematical transformations in order to yield scores with a range of approximately 0–5 and a central mode. Transformations were applied equivalently within each measure across all three waves of data. Counts of prior child welfare referrals were log transformed for analyses and truncated at 15 for graphical presentations. Multiple clusterings and cross-classifications were present in the data, both overt and covert (e.g., waves nested within subjects, within home visitors, within teams, within agencies, cross-classified by county offices, and cross-classified with adjunctive service provider agencies). Given that no effects of interest were being modeled at levels beyond subjects, that some clusterings and cross-classifications were covert or unobserved, and given the limits of model complexity that can practically be accommodated, we opted to employ robust maximum likelihood estimation and a sandwich estimator to manage possible cluster dependencies by adjusting standard errors (Muthén & Asparouhov, 2002). Latent difference score (LDS) models were used to represent change over time. Readers interested in a fuller description of the LDS approach should consult McArdle (2009) or King et al. (2006). Difference scores are useful in examining point-to-point change, but simple or unadjusted difference scores suffer from a number of limitations. Simple difference scores are vulnerable to regression toward the mean, ceiling or floor effects, and poor reliability due to autocorrelation or slope–intercept relationships. The LDS change model allows us to estimate true change over time while accounting for these potential sources of variation. Figure 1 shows the LDS structural model used in this study. For the sake of presentation simplicity, response bias corrections (i.e., Child Abuse Prevention Inventory Lie Scale paths) are not displayed in the figure but were included in the model. The model assumes that the observed scores (CAPI, BDI, FRS, and SPS) at each time point reflect a common latent factor (the Iy 's), plus independent elements not reflected in the common latent factor and measurement error (the ψ 's). These true risk scores (the Iy 's) are not directly observed but can be inferred. This is the measurement model. Factorial invariance was obtained using equality constraints. The latent intercept ($Iy1$), and the two LDS ($Iy1$ and $Iy2$) describe each individual's change trajectory and can be treated as random variables for a single parametric trajectory or can be predicted nonparametrically by a latent class variable to identify change trajectory classes in a mixture model.

The approach used to test child welfare system recidivism survival was a discrete-time recurrent event survival model. The recurrent event approach was selected in order to better capture the link to future chronicity, given that chronicity implies multiple recurrent events rather than simply a first event as analyzed in basic survival models. Recidivism data, drawn from child welfare administrative databases, were available for all study participants. The gap time coding and latent variable modeling framework described by Masyn (2009) was employed for the recurrent event analysis. Sequential 90-day discrete intervals were coded for presence or absence of a new child welfare report. If an event occurred or the end of administrative follow-up was reached, remaining intervals in the sequence were censored. In the gap time coding scheme, each new event resets the event clock to zero for a new subsequent interval sequence. These sequences were modeled as reflecting a single latent

variable that could be either random (i.e., a parametric frailty model) or predicted by a latent class variable (i.e., a nonparametric model).

Missing data—There were no large observed selection effects with respect to missing data patterns. Missing data pattern groupings (three waves, two waves, and one wave) were not significantly bivariately related to the number of previous child welfare reports or to baseline household income, parent education, the number of children in the family, or scores on the Child Abuse Prevention Inventory, BDI, and SPS measures. Missing data pattern had a statistically significant but practically negligible relationship with baseline FRS score (means between 3.8 and 3.9, $SD = 0.57$; Eta Squared = 0.003, $p = .04$) and parent age (means between 29 and 30 years, $SD = 8.0$; Eta Squared = 0.006, $p = .04$). Missing self-report data patterns were not prospectively related to recidivism hazard drawn from administrative data. Missingness was managed using robust maximum likelihood estimation under the covariate dependent missing at random assumption.

Results

Confirmatory Testing of the Measurement Model

Confirmatory testing of the measurement model was undertaken in order to examine whether a single latent factor was warranted for the observed variables (i.e., Child Abuse Prevention Inventory, BDI, FRS, and SPS) and whether its structure was stable over time. Beginning with baseline data, all loadings for the Child Abuse Prevention Inventory, FRS, SPS, and BDI on the single latent variable in the confirmatory factor analysis were highly significant (Estimate/ SE from 23.8 to 49.9, all $p < .0001$). Omega for the factor was .90 and the CFI was .97. In order to examine the desired assumption of factorial invariance over time, a second confirmatory factor analysis was conducted creating three separate latent variables, one for each of the three waves, allowing the latent variables to correlate, fixing latent factor variances to one, and leaving observed variable loadings, intercepts and residuals unconstrained. Factor loadings, intercepts, and residuals across the three waves were very close in value, and imposing equality constraints did not significantly alter model fit (Wald = 9.0, $df = 8$, $p = .34$), supporting factorial invariance over time.

Fit of a Single Parametric Trajectory Model

The basic LDS model fit the observed data modestly well (Standardized Root Mean Square Residual [SRMR] = 0.08; root mean square error of approximation [RMSEA] = 0.09; Comparative Fit Index [CFI] = 0.82), and residual variances for the two LDS were large and highly significant, implying that a single parametric change trajectory may fit this data only modestly well and that unobserved heterogeneity in change trajectories might be present. This finding supports moving forward with a nonparametric class approach to modeling trajectories.

Latent Class Model of Trajectories

Next, a latent class predictor was incorporated into the LDS model as shown in Figure 1. In this mixture model, the heterogeneous classes of interest represent different trajectories as defined by $ly1$, $ly1$, and $ly2$ the in the LDS model. For an overview of modeling

heterogeneity in change trajectories, see Nagin and Odgers (2010). Means and lagged proportional effects were allowed to be class-specific (Lubke & Muthén, 2007). There is no single formula for settling on the number of latent classes extracted by mixture models. BIC scores, entropy scores which reflect classification confidence, trajectory distinctiveness, and interpretability of the classes were all considered in this process. One, two, three, four, five, six, and seven class models were sequentially estimated before settling on the five-class model as offering the best BIC (the six class model fit was only very slightly better), entropy, and distinctiveness combination. Additional trajectories extracted in the six and seven class models largely duplicated existing classes in the five class model. The BIC for the five-class mixture model was 51,677, representing an improvement in fit over the value of 51,909 for the single parametric trajectory LDS model, and residual variances for the LDS were substantially smaller and no longer significant for $ly2$. Because the posttreatment and follow-up data were collected at benchmark points that varied somewhat in time, we tested whether either of the two difference scores estimated by the model was affected by its corresponding interwave time interval in days. Neither correlation ($r = -.01$ and $r = .02$) was significant.

The five model trajectories are shown in Figure 2. For descriptive purposes and following the hypotheses described in the introduction, the trajectories were labeled as Stable Low, Relapsing, Sustained Improvement, Stable High, and Paradoxical. Although details are not reported here, the five class model was reexecuted using listwise deletion (i.e., only cases with no missing data), and essentially duplicated the trajectory pattern that emerged from the full data set, although with the expected higher level of classification certainty. Class membership probabilities obtained from the model using all participants were correlated with their listwise deletion model counterparts $r = .94$ for the Stable Low class; $r = .81$ for the Relapsing class; $r = .76$ for the Sustained Improvement class; $r = .66$ for the Stable High class; and $r = .42$ for the Paradoxical class. Missing data pattern was bivariate unrelated to class membership probabilities for the first four classes but was slightly related to class membership probability for the Paradoxical class (Eta Squared = 0.006, $p < .01$). Together, these findings suggest that the Paradoxical class may have been impacted more by missing data.

In order to examine the construct validity of the classes, class membership was tested for association with independently collected change indicators drawn from administrative clinical record data. Home visitors recorded a posttreatment progress rating indicating the parent's overall level of risk related goal attainment (none, some, most, or all). These ratings were recorded by clinicians as part of routine record keeping and were made independent of the client self-report data that were used to create the trajectories. Note that because home visitor goal attainment ratings were made only at posttreatment, we would expect them to correspond with trajectory midpoints, not follow-up endpoints. We would predict that the trajectories with lower problem levels at post-treatment would be rated by clinicians as having better goal attainment than the trajectories with higher problems at posttreatment. The correspondence between trajectory membership and clinician ratings was evaluated using an ordinal logistic model, which demonstrated a significant overall association (Likelihood Ratio Chi-Square = 31, $df = 4$, $p < .001$, E.S. = .39). The rank order of model predicted goal attainment ratings (from best to worst) was Stable Low, Sustained

Improvement, Relapsing, Paradoxical, and Stable High, which is consistent with the rank order of post-treatment problems that can be seen in the trajectories (see Figure 2). Demographic data for participants classified into each of the five trajectories are presented in Table 1, which also display baseline values on the variables used to form the trajectories and demonstrate how differing trajectory intercepts reflect substantial differences in baseline measures of risk used to form the trajectories.

Change Trajectory Classes and Chronicity

A multinomial logistic regression model was executed to predict class membership from the log transformed number of prior child welfare referrals. The Stable Low trajectory class, which had the most members (see Figure 2) was set as the reference class. The number of past referrals significantly predicted class membership overall (Chi-Square = 27.7, $df = 4$, $p < .001$). More prior child welfare referrals was associated with increasing probability of membership in the Sustained Improvement class ($\text{Exp}(\beta) = 1.45$, 95% CI = [1.14, 1.84], $p < .01$) and the Stable High class ($\text{Exp}(\beta) = 1.42$, 95% CI = [1.21, 1.67], $p < .001$), relative to the Stable Low class. Relapsing and Paradoxical class membership had no significant relationship with prior referral count. Predicted probability of class membership as a function of prior referral count is shown in Figure 3.

Recidivism and Change Trajectory Classes

Recidivism events were recorded beginning at baseline. The average follow-up period for recidivism events was 1,048 days, which extended on average around 600 days beyond the last measurement point for self-report variables but also included intervals between the self-report measurement points. This raises the issue of whether change trajectories derived from the self-report variables may be temporally confounded with recidivism. Recidivism prior to measurement might influence the change trajectory and this influence might differ across trajectory classes. This would be reflected in some change trajectories having a different proportion of events occurring prior to, rather than after, post-baseline measurement points. As a check on this, we pretested the trajectory classes for differences in the timing of first recidivism events vis a vis the timing of post-treatment and follow-up measurement waves. No significant differences among trajectory classes were found.

About 40% of all participants had one or more future reports, 24% had two or more, 13% had three or more, 7% had four or more, and 4% had five or more, which translated into up to five discrete interval sequences that potentially could be populated for any individual participant. In the recurrent event survival model, the parametric frailty model and the nonparametric model fit the data equally well, and the parametric frailty variable was used for subsequent analyses. Class membership significantly predicted recidivism hazard. Relative to the Stable Low class, recidivism was significantly higher for the Relapsing class (Estimate = .22, $\text{Est}/SE = 3.7$, $p < .001$; Effect size = 0.41), the Stable High class (Estimate = .23, $\text{Est}/SE = 8.09$, $p < .001$; Effect size = 0.43), and the Paradoxical class (Estimate = .23, $\text{Est}/SE = 4.33$, $p < .0001$; Effect size = 0.43). The Sustained Improvement class had a smaller but still statistically significant difference from the Stable Low class (Estimate = .13, $\text{Est}/SE = 3.19$, $p < .01$, Effect size = 0.24). Setting the Stable High class as the reference class, the Relapsing and Paradoxical classes did not significantly differ, but the Sustained

Improvement class had a significantly lower recidivism hazard (Estimate = -0.11 , $Est/SE = -2.42$, $p < .05$, Effect size = -0.20). The next step was to test whether the relationship between change trajectory class and recidivism was moderated by the number of prior referrals. Adding the count of prior child welfare referrals to the model, the number of prior referrals was a significant predictor of recidivism hazard, as expected (Estimate = 0.05 , $Est./SE = 9.86$, $p < .001$). Change trajectory class membership remained significant with only slightly reduced estimates and critical ratios from those reported immediately above. Prior referral count did not significantly moderate the relationship between change trajectory class and recidivism, evaluated both via class specific interaction terms in the model and via comparing free versus equality constrained multigroup models. Recidivism risk as a function of class membership and prior reports are displayed in Figure 4. The figure depicts recidivism effect sizes (i.e., relative risks) across trajectory class membership combined with chronicity level and summarizes a number of overall findings, including the relatively low risk associated with Sustained Improvement even at high levels of chronicity, and the high relative risk associated with Relapsing and Stable High trajectories, especially when combined with high chronicity. Because the study population was high risk in general, the relative risks observed (i.e., over 2.0 and approaching 2.5) suggest very serious risk for parents in these trajectories.

Discussion

Some of the predications about change trajectories and the relationship between case chronicity and change trajectories were supported, but others were not. Perhaps, the most strongly supported prediction was the idea that chronic cases enter services with high problem levels and often show limited and probably insufficient improvement over the course of services. The odds of being in the Stable Low class decreased and the odds of being in the Stable High class increased as number of prior reports increased. At the margins of high chronicity, the Stable High trajectory became the modal change trajectory class. And when a Stable High pattern was observed, recidivism was high. This is consistent with predictions of the “untreatable family” hypothesis. Jones (1987) questioned the wisdom of devoting intensive services to these types of cases, given that little is likely to be gained and the service resources might be more productively targeted toward better prognosis cases. Alternately, one might argue that better change might have been observed, given a different type or quality of services, but there was little evidence available to either support or refute this.

But the untreatable family hypothesis was not supported in all respects and circumstances appeared to be more complex than this hypothesis might suggest. Most notably, chronicity also was associated with a greater probability, not a lesser probability, of falling into the Sustained Improvement trajectory class. This is an encouraging finding with respect to chronic cases, suggesting that they retain a reasonable chance of meaningful improvement during services. Cases in this trajectory had high problem levels, showed good improvement over the course of services, their improvement was sustained at follow-up, and membership in the trajectory was associated with lower downstream recidivism. This favorable trajectory was noted in 10–20% of all cases. The modest odds might suggest that attaining sustained improvement among high problem cases can require multiple service episodes, similar to

patterns that have been observed in the drug dependency literature (Hser et al., 1998). The increasing chances of Sustained Improvement with increasing chronicity might suggest a cumulative dose phenomenon. Alternately, this finding could be an artifact of the study design. Because the study relied on a cross-sectional cohort, it is possible that poorer prognosis chronic parents eventually lost custody of their children, and thus were progressively pruned from the sample over time. This might result in longitudinal selection bias, favoring elimination of the worst prognosis cases among higher chronicity groups.

Other possible relationships between chronicity and change during services were not supported. Little support was observed for the idea that many clients get worse over the course of services and that this is associated with chronicity. The relapsing hypothesis, which suggests that chronic clients improve substantially during services but worsen afterward, also received little support. The Relapsing Class and the Paradoxical Class were both observed infrequently and class membership had little relationship to accumulated prior reports. The Paradoxical Class appeared potentially vulnerable to missing data artifacts, so findings surrounding this class should be viewed cautiously. The Relapsing Class did appear particularly concerning in one key respect—it showed the highest recidivism of any class. One concerning element of this observation is that Relapsing cases were indistinguishable from Sustained Improvement cases on observed measures at the posttreatment assessment point. In other words, observing substantial improvement from a high problem baseline may convey distinct prognoses depending on whether that progress is sustained or not. This key distinction would be difficult or impossible to make, given the episodic nature of child welfare services and the lack of follow-up after case closure. Consequently, predicting prognosis on the basis of baseline to posttreatment improvement among high problem cases may not be possible—the prognosis could be quite favorable or quite poor depending on whether gains are sustained.

The findings might raise a number of questions about service system structure. We would suggest that the episodic and reactive service model characterizing traditional child welfare services may be a mismatch with the trajectories increasingly observed among families who are chronically in the system. Episodic service models presume that once a condition (i.e., maltreatment) is detected, a rehabilitative treatment for that condition can be applied and then the treatment is done. Services are only reinstituted if the same or another problem recurs, and in this sense the model is reactive rather than proactive over time. Episodic and reactive service models are better suited to acute conditions but are a mismatch with chronic conditions. Chronic care models more often rely on proactive, lower intensity, longer term approaches that emphasize monitoring, maintenance, stepped care, or harm reduction. Chronic care models may be especially suited to “untreatable family” cases (e.g., harm reduction) or to cases that may eventually or cumulatively respond to services (e.g., extended or stepped-care models) or to cases where the durability of change needs to be ascertained past the point of case closure (e.g., monitoring models). None of these alternative chronic care approaches are a particularly good fit within the child welfare system as it is currently constructed, although they may be better suited to the types of cases increasingly being seen in child welfare (Kaplan et al., 2009).

Some limitations of the study need to be considered. Foremost among these is the reliance on a single cross-sectional cohort. We cannot determine from this data whether the observed trajectories are stable across time (as the untreatable family hypothesis might suggest) or evolve (as the cumulative benefit hypothesis might suggest). Although recidivism was prospectively ascertained, chronicity was retrospectively ascertained and therefore may have been vulnerable to longitudinal attrition bias. Studies of cumulative treatment benefit or treatment resistance would benefit from fully prospective designs that begin with a cohort of first time child welfare cases, follow them forward over time, and examine change during each successive service episode. This analysis examined correlations and associations within subjects over time and was not intended to test whether any type of home-based services are or not effective. The change trajectories observed may or may not have been directly caused by the services provided. Data were drawn from a single statewide system, so generalization to other contexts needs to be explored. This study also relied on self-report measures of risk factors. Although response bias correction procedures were employed, other source effects may have been operating. A composite latent factor indicator of risk was used, and although this factor reflected commonality across a range of domains (parenting and actuarial risk, mental health, basic needs, and social support), all relevant change dimensions may not be reflected by this factor. Five trajectory classes were extracted, but mixture modeling may not yield stable or genuinely taxonomic classes, so the trajectories extracted here should not be taken as absolute. Finally, the amount of missing data in the study was substantial and although we did not observe evidence of bias, the presence of unobserved bias cannot be ruled out.

In summary, the findings suggest that many high problem chronic cases in child welfare show limited change with services. However, this finding is tempered by the fact that sustained improvement still remains possible and that when this does occur, recidivism risk is reduced even among the most chronic cases in the system. This is consistent with recent administrative data findings also indicating persistent home-based service benefits across levels of chronicity (Jonson-Reid, Emery, Drake, & Stahlschmidt, 2010). Thus, the prognosis for these cases, and the wisdom of trying the same or very similar services again, is complex. Although reactive episodic service efforts are not necessarily wasted on chronic families, chronic care models ought to be considered as a potentially better fit for this growing segment of the child welfare population.

Acknowledgments

The authors wish to recognize the contributions of John Lutzker; Randy Campbell; Kathy Bigelow; Jill Filene; Dan Whitaker; Greg Aarons; Steve Ross; Gina Carrier; staff and leadership at the Oklahoma Department of Human Services including Howard Hendrick, B. K. Kubiak, J. J. Jones, John Gelona and Kathy Simms; and the leadership and staff of the Oklahoma Children's Services network agencies

Funding

The author(s) disclosed receipt of the following financial support for the research and/or authorship of this article: This project was supported by grant number R01MH065667 from the National Institute for Mental Health. Additional in kind support was provided by the Violence Prevention Branch of the U.S. Centers for Disease Control and Prevention..

References

- Beck, AT., Steer, RA., Brown, GK. BDI-II Manual. San Antonio, TX: The Psychological Corporation; 1996.
- Beck AT, Steer RA, Ball R, Ranieri WF. Comparison of beck depression inventories-IA and -II in psychiatric outpatients. *Journal of Personality Assessment*. 1996; 67:588–597. [PubMed: 8991972]
- Chaffin M, Kelleher K, Hollenberg J. Onset of physical abuse and neglect: psychiatric, substance abuse, and social risk factors from prospective community data. *Child Abuse & Neglect*. 1996; 2:191–203.
- Chaffin M, Valle LA. Dynamic prediction characteristics of the child abuse potential inventory. *Child Abuse & Neglect*. 2003; 27:463–481. [PubMed: 12718957]
- Cutrona, CE., Russell, D. The provision of social relationships and adaptation to stress. In: Jones, WH., Perlman, D., editors. *Advances in personal relationships*. Vol. 1. Greenwich, CT: JAI Press; 1987. p. 37-67.
- DePanfilis D, Zuravin SJ. Rates, patterns, and frequency of child maltreatment recurrences among families known to CPS. *Child Maltreatment*. 1998; 3:27–42.
- DePanfilis D, Zuravin SJ. Predicting child maltreatment recurrences during treatment. *Child Abuse & Neglect*. 1999; 23:729–743. [PubMed: 10477234]
- Drake B, Jonson-Reid M, Sapokaite L. Rereporting of child maltreatment: does participation in other public sector services moderate the likelihood of a second maltreatment report? *Child Abuse & Neglect*. 2006; 11:1201–1226.
- Drake B, Zuravin S. Bias in child maltreatment reporting: Revisiting the myth of classlessness. *American Journal of Orthopsychiatry*. 1998; 68:295–304. [PubMed: 9589767]
- Dunst CJ, Leet HE. Measuring the adequacy of resources in households with young children. *Child: Care, Health, and Development*. 1987; 13:11–125.
- English DJ, Marshall DB, Orme M. Characteristics of repeated referrals to child protective services in Washington state. *Child Maltreatment*. 1999; 4:297–307.
- English DJ, Upadhyaya MP, Litrownik AJ, Marshall JM, Runyan DK, Graham JC, Dubowitz H. Maltreatment's wake: The relationship of maltreatment dimensions to child outcomes. *Child Abuse and Neglect*. 2005; 29:597–619. [PubMed: 15970327]
- Ethier LS, Lemelin J, Lacharite C. A longitudinal study of the effects of chronic maltreatment on children's behavioral and emotional problems. *Child Abuse and Neglect*. 2004; 28:1265–1278. [PubMed: 15607769]
- Finkelhor, D., Jones, L., Shattuck, A. Updated trends in child maltreatment, 2008. Durham, NH: Crimes against Children Research Center; 2010.
- Fluke J, Yuan Y, Edwards M. Recurrence of maltreatment: An application of the national child abuse and neglect data system (NCANDS). *Child Abuse and Neglect*. 1999; 23:633–650. [PubMed: 10442829]
- Graham JC, English DJ, Litrownik AJ, Thompson R, Briggs EC, Bangdiwala SI. Maltreatment chronicity defined with reference to development: Extension of the social adaptation outcomes findings to peer relations. *Journal of Family Violence*. 2010; 25:311–324.
- Hecht, D., Chaffin, M., Silovsky, JF. Annual evaluator's report on Oklahoma Comprehensive Home Based Services. Oklahoma: University of Oklahoma Health Sciences Center; 2010.
- Hser Y, Grella C, Chou C, Anglin MD. Relationships between drug treatment careers and outcomes: Findings from the National Drug Abuse Treatment Outcome Study. *Evaluation Review*. 1998; 22:496–519.
- Jones DP. The untreatable family. *Child Abuse and Neglect*. 1987; 11:409–420. [PubMed: 3315132]
- Jonson-Reid M, Drake B, Chung S, Way I. Cross-type recidivism among child maltreatment victims and perpetrators. *Child Abuse & Neglect*. 2003; 27:899–917. [PubMed: 12951139]
- Jonson-Reid M, Emery CR, Drake B, Stahlschmidt. Understanding chronically reported families. *Child Maltreatment*. 2010; 15:271–281. [PubMed: 20941889]
- Kaplan C, Schene P, DePanfilis D, Gilmore D. Shining light on chronic neglect. *Protecting Children*. 2009; 24:1–7.

- King LA, King DW, McArdle JJ, Saxe GN, Doron-LaMarca S, Orazem RJ. Latent difference score approach to longitudinal trauma research. *Journal of Traumatic Stress*. 2006; 19:771–785. [PubMed: 17195976]
- Lemmon JH. The effects of maltreatment recurrence and Child Welfare services on dimensions of delinquency. *Criminal Justice Review*. 2006; 31:5–32.
- Lilienfeld SO. Psychological treatments that cause harm. *Perspectives on Psychological Science*. 2007; 2:53–70. [PubMed: 26151919]
- Lubke G, Muthen BO. Performance of factor mixture models as a function of model size, covariate effects and class specific parameters. *Structural Equation Modeling*. 2007; 14:26–47.
- Mancini JA, Blieszner R. Social provisions in adulthood: Concept and measurement in close relationships. *Journal of Gerontology*. 1992; 47:14–20.
- Masyn KE. Discrete-time survival factor mixture analysis for low-frequency recurrent event histories. *Research in Human Development*. 2009; 6:165–194. [PubMed: 24489519]
- McArdle JJ. Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*. 2009; 60:577–605.
- McLellan, AT., Lewis, D., O'Brien, CP., Hoffman, NG., Kleber, HD. Is drug dependence a chronic medical illness?. Philadelphia, PA: Treatment Research Institute; 2000. Retrieved September 28, 2010, from <http://www.lhc.ca.gov/studies/activestudies/drug/McClellanApr25.pdf>
- Milner, JS. The child abuse potential inventory manual. 2. DeKalb, IL: Psytec; 1986.
- Muthén, BO., Asparouhov, T. Using Mplus Monte Carlo simulations in practice: A note on non-normal missing data in latent variable models. 2002. Retrieved October 5, 2009, from www.statmodel.com/examples/webnote.shtml
- Muthén, BO., Muthén, LK. Mplus user's guide. 6. Los Angeles, CA: Author; 2010.
- Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. *Annual Review of Clinical Psychology*. 2010; 6:109–138.
- Nelson KE, Saunders EJ, Landsman MJ. Chronic child neglect in perspective. *Social Work*. 1993; 38:661–671. [PubMed: 8256141]
- U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau. Child Maltreatment 2008. 2010. Available from http://www.acf.hhs.gov/programs/cb/stats_research/index.htm#can
- Widom, CS. Long-term consequences. Presentation at Chronic Neglect Research Meeting of the American Humane Association; Denver, CO. 2009 Jan.

Figure 1.

Structural model for trajectories and trajectory classes. *Notes:* BDI = Beck Depression Inventory; CAPI = Child Abuse Potential Inventory; FRS = Family Resources Scale; SPS = Social Provisions Scale. Numbers 1–3 correspond to wave. Additional paths are not depicted for response bias corrections at each wave using the CAPI Lie Scale for all observed variables.

Figure 2.
Change trajectory classes.

Figure 3.
Probability of change trajectory class as a function of prior referrals.

Figure 4.
Relative recidivism risk as a function of past referrals and change trajectory class.

Table 1**Baseline Comparison of Change Trajectory Classes**

Trajectory Class	Stable Low	Relapsing	Sustained Improvement	Stable High	Paradoxical
Demographic characteristics					
Female (%) *	89	91	94	94	97
Non-Hispanic Caucasian (%)	67	62	73	65	69
Married (%) *	32	40	39	28	21
Less than H.S. educ. (%) *	38	39	38	43	48
Urban residence (%)	31	21	20	25	27
Women currently pregnant (%)	8	8	6	7	11
Age (<i>M</i>)	29	29	29	29	30
Age at birth of first child (<i>M</i>)	22	22	22	22	23
Number of children (<i>M</i>)	2.8	2.6	3.0	3.0	2.9
Monthly household income \$ *	1,287	1,255	1,041	1,045	903
Measured risk factors					
Beck Depression Inventory Score (<i>M</i>) *	5.9	12.4	20.3	25.9	9.7
Significant depression (BDI > 19 (%)) *	2	21	56	73	27
Child Abuse Prevention Inventory Abuse Score *	88	147	247	294	124
Child Abuse Prevention Inventory Lie Score (<i>M</i>) *	8.4	9.8	5.9	5.9	9.0
Family Resources Scale (<i>M</i>) *	4.0	3.9	3.6	3.5	3.9
Social Provisions Scale (<i>M</i>) *	3.4	3.1	2.9	2.8	3.2
Follow-up times					
Baseline to Posttreatment days (<i>M</i>)	227	213	246	232	231
Post to follow-up days (<i>M</i>)	221	236	204	216	219
Recidivism follow-up days (<i>M</i>)	1,033	1,065	1,073	1,062	1,048

*Note.** $p < .05$ by analysis of variance (ANOVA) for means and chi-Square for percentages.