PARTNERS IN CARE 6

Risk and Resilience

Ontario Association of Residences Treating Youth

Centre of Excellence for Residential Care in Ontario

Table of Contents

Risk and Resilience
Partners In Care 6

1.0 Background

The Ontario Association of Residences Treating Youth (OARTY) maintains a data repository of anonymized clinical information on the clients served by member agencies. The repository contains clinical profiles on 4,616 unique clients from 98 agencies. A total of 1,361 new clients have been added this year for PIC 6. Please see the appendix for the repository details.

The research hypothesis of this study was that the children placed in the last four years would show significant growth in resilience and the resilience would buffer these children against the adverse consequences of risk in their lives.

1.1 Synopsis of Findings

A critical mass of OARTY members representing 25% of all children served have implemented a system of outcome measurement where each child is tested multiple times on the same set of instruments. Analyzing this dataset, OARTY has been able to report:

a) We found a difference between the children placed in Family Based Care (FBC) compared to Staff Operated Group Care (Group). Children in FBC and Group vary by their clinical profile[1]; age at the time of placement and amount of adversity and trauma. The children in Group are 4 years older and more disturbed.

b) Children in both Group and FBC are improving to a large degree in the 1st year of service in areas such as positives, functioning and risk

c) Children are at great risk of mental illness (59%) because they display high risk symptoms. The children make a large improvement in reducing the risk. Nineteen percent (19%) of the children who were at-risk when 1st tested show zero risk in less than one year of service.

d) The research hypothesis about resilience and risk was confirmed by the study.

---

[1] Clinical profile includes the type of child by diagnosis, nursing care needs and complex neurological disorders.

2.0   Research Design: 2015

The OARTY data repository has collected a snapshot across different time periods of the adversity and clinical profile of the children placed in Group and FBC in Ontario. The data repository follows 2,493 children with results of outcome testing; some were repeated up to 13 times. The outcomes measured are positives, functioning, the risk of mental illness and the degree of impairment. The research questions are:

1) What is the clinical profile of the children in different settings and at different times?
2) What is the pattern of placement in family based care and staffed group care?
3) What are the outcomes for different groups of children?
4) What are the outcomes for children in family based care compared with children in staffed group care?
5) Do children acquire greater resilience over time?
6) Is there a relationship between resilience, functioning, risk and impairments?

2.1   Subject of the PIC 6 Research Study

Resilience and Risk

A literature review of the nature of resilience in the Journal of Psychiatry offered these definitions of resilience[2]:

1) the ability to maintain or regain mental health, despite experiencing adversity;
2) the protective factors and processes or mechanisms that contribute to a good outcome, despite experiences with stressors shown to carry significant risk for developing psychopathology; and
3) a multi-dimensional characteristic that varies with context, time, age, gender and cultural origin, as well as within an individual subject to different life circumstances.

Measurement of Resilience

Afifi & MacMillan (2011) examined ways of measuring resilience. They found that resilience is "frequently measured using normative levels of internalizing and externalizing symptoms and competent age-appropriate functioning across several developmental domains.. Using multiple indicators to determine resilience is a widely accepted practice a person may show competence in one domain, but not in another".[3]

The strategy used by OARTY to measure resilience is frequently cited in the literature as noted by Afifi & MacMillan (2011). Two composite variables were created to measure resilience: positives and functioning.

---

[2] Herrman, Helen; Stewart, Donna E; Diaz-Granados, Natalia; Berger Elena L; Jackson, Beth & Yuen, Tracy (2011), "What is Resilience?", *Canadian Journal of Psychiatry*, 56(5): pp 259, 260

[3] Afifi, Tracie O & MacMillan, Harriet L (2011), "Resilience Following Child Maltreatment: A Review of Protective Factors", *Canadian Journal of Psychiatry*, 56(5), p. 267
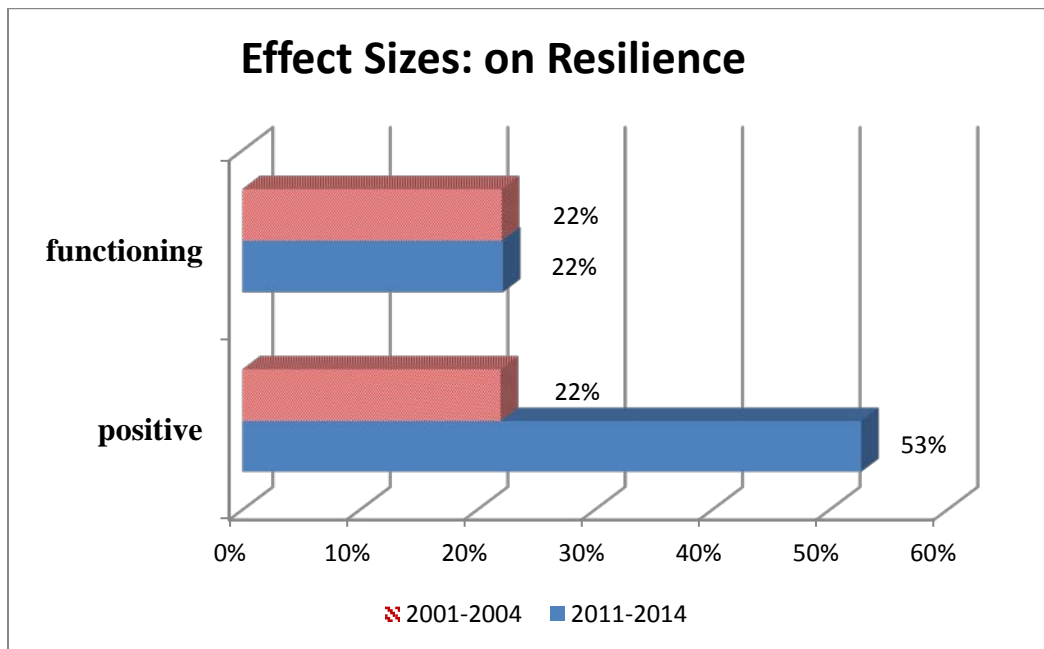
OARTY members, participating in this study, used combinations of nine instruments to measure qualities of strengths and positives. These instruments measure attachment, pro-social behavior, happiness, peer acceptance, quality of life, self-efficacy and social support. The effect sizes for each instrument were calculated and a composite score of the direction and degree of change in the construct, *positives*, was computed by the average Cohen's d for the component tests.

OARTY members used combinations of eight instruments to measure qualities that are indicative of functioning. These instruments measure academic performance, social role performance, independence, adaptive skills, self-esteem, self-efficacy and social skills. The effect sizes for each instrument and the average Cohen's d for the component tests were calculated.

3.0   Building Strengths/Growing Positives

Children in the care of OARTY member agencies are making large improvements in the growth of positives in their first year of service. The improvement is continuing for many years. In this section, we will contrast the rate of improvement for males and females, family based care and staffed group care, children by cultural groups, and children by different diagnostic categories, including FASD, autism, other complex neurological disorders, psychiatric disorders and symptom free. We will study whether the changes in positives are stable across the decade.

3.1   Resilience: Findings

Effect Sizes: on Resilience

- functioning: 22% (2001-2004), 22% (2011-2014)
- positive: 22% (2001-2004), 53% (2011-2014)

Legend: 2001-2004, 2011-2014

Children make steady improvements in functioning and positives. In the last four years the growth in positives has been large. The treatment effects observed in the years 2011 to 2014 (Cohort 2) were compared with the treatment effects obtained for children placed in 2001-2004 (Cohort 1).

3.2   Positives

The children from Cohort 2 produced a 53% improvement in positives. The graph below follows a high risk group of children across 36 months of treatment. Forty-one percent (41%) of children start treatment at very high risk on an attachment scale (PBI). They have on average *caring scale* t-score of 31 on their first test. Less than 1 in a 1,000 children across the English speaking world produce a score this low. These children feel a profound lack of affection in their lives.
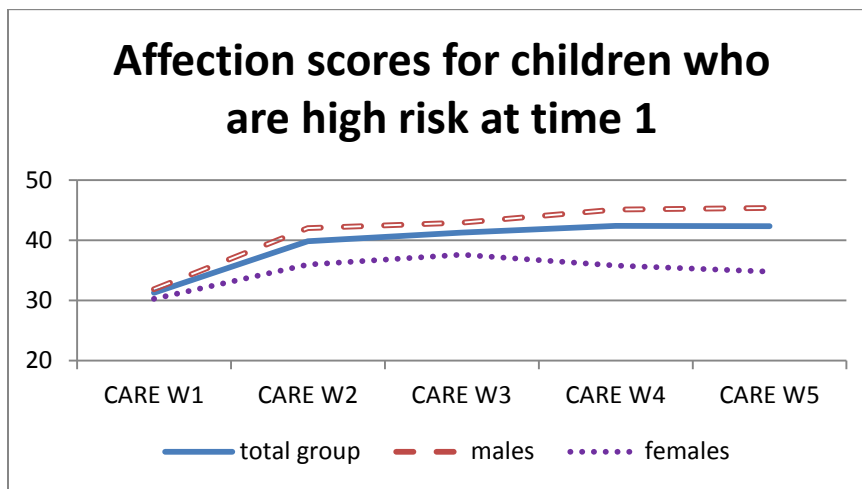
All children, indicated by the solid line, improved by 86% by the 2nd test. Males made larger gains throughout the path and females didn't finish as well as males after making a large initial improvement.

Children from cohort 2 made greater improvements on measures of positives (53%) compared to children in the past (22%). Multifactor analysis of variance was used to determine what difference is driving the change. These are the findings:

(1) Factors internal to the agencies contribute more to the different outcomes than the cohort: *cohort effects*: F-ratio = 4.271, p = .04; *agency effects*: F-ratio = 6.731, p = .01.

(2) Differences in the mix of children in different cohorts contribute more to the different outcomes than type of placement (i.e.) FBC vs Group: *FBC effects*: F-ratio = 1.854, p = .174; *cohort effects*: F-ratio = 7.657, p = .006.

   a. There is a significant interaction between outcomes due to the cohort with and the type of placement in each cohort sub-group (decade vs recent): *interaction effect*: F-ratio = 4.12, p = .043

(3) Factors internal to the agencies contribute more to the different outcomes than the type of placement (i.e.) FBC vs Group: *FBC effects*: F-ratio = 0.858, p = .355; *agency effects*: F-ratio = 9.566, p = .005.

Ten agencies measured positives multiple times; 60% of these agencies provided both FBC and Group. There were statistically significant differences between the agencies in the amount of change each produced on positives. Each program offering FBC had a different result and so did the programs offering Group. Multifactor analysis of variance computes how much of the difference in outcomes could be attributed to the agency level compared to the program level. The average outcome on positives after the 1st year by agency explains the variance by a factor of 9.5 to one.

When you hold the agency constant, the change that each child makes comparing FBC and Group shows no difference.

There is an *interaction effect* between the cohort and the type of placement. Specifically:

a) Children placed in FBC recently (Cohort 2) show better outcomes in terms of positives compared to Group (**FBC** m = 80%, **Group** m = 36%, , diff = 44%, t = 2.017, n = 161, p = .047)

b) This advantage in outcomes for FBC disappears for children placed in 2001-2004.

As noted above, the response of individual children placed in each agency, in FBC or Group, varied by agency not by the type of placement within the agency. Multifactor analysis found that the individual response of children is affected by when the service happened. There is an apparent advantage to children placed in FBC recently but this was not the case a decade earlier. This is referred to as an interaction effect. This is not a main effect; the major finding is true regardless of when the service occurs.

The interaction effect means that, within each agency, individual children placed in FBC are responding better in terms of positives in the recent period compared to a decade earlier. This type of interaction can occur if

- Agencies have changed their practices delivering FBC
- Agencies are providing more resources to children placed in FBC
- The clinical profile of the children placed in FBC has changed across the decade, and, if
 - The changes in the clinical profile mean better outcomes

The clinical profile in FBC in Cohort 1 compared to Cohort 2 has changed as indicated in the next section.

Changing Profile of Children in Family Based Care

There has been a 20% shift in the types of placements selected across the decade from Group to FBC. However, some groups of children have seen a drop in the percentage of placements in FBC and other groups have seen the use of FBC increase.

Thirty one percent (31%) of all children served are in one of these high risk groups.
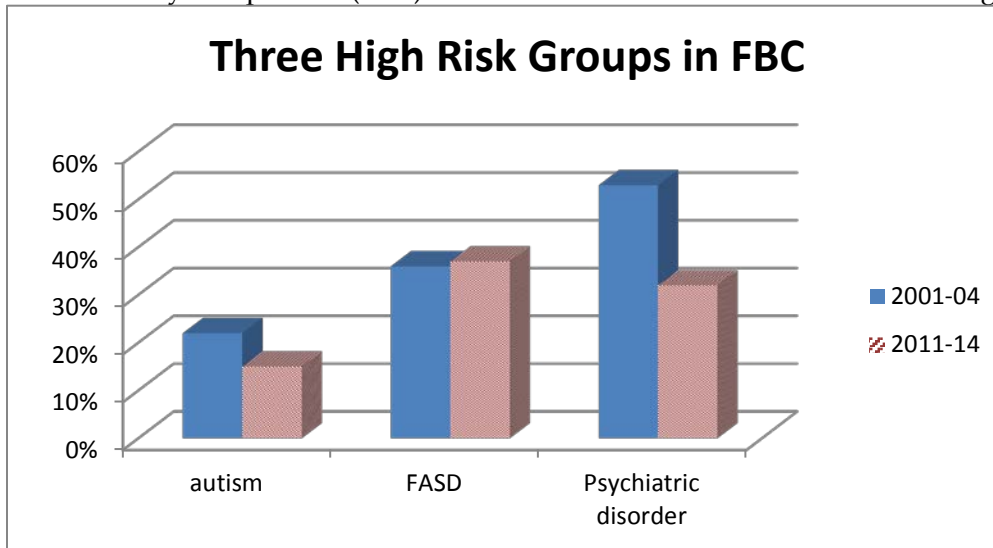
The proportion of children with a high risk diagnoses who are placed in FBC has decreased across the decade: autism (22% to 15%) and psychiatric diagnoses (53% to 32%). The percentage of children with FASD in FBC has remained stable (36% to 37%).

Group has always been the preferred placement for high risk children and this pattern is accelerating. The finding that high risk children are preferentially placed in staffed group care has been reported many times in the literature. A study by the Child Welfare League of America (2005) followed 2,100 children placed in staffed group care and treatment foster care for four years. They found that children in both settings were very disturbed on admission, but that staffed group care had much greater frequencies of children severe mental illness, delinquency and school problems. [4]

---

[4] *The Odyssey Project: A Descriptive and Prospective Study of Children and Youth in Residential Group Care and Therapeutic Foster Care* (2005), Alicia A. Drais-Parrillo, Ph.D. (ed), Child Welfare League of America: Washington, DC, page 22

Twelve percent (12%) of all children served are in one of the following low risk groups.

Three Low Risk Groups in FBC
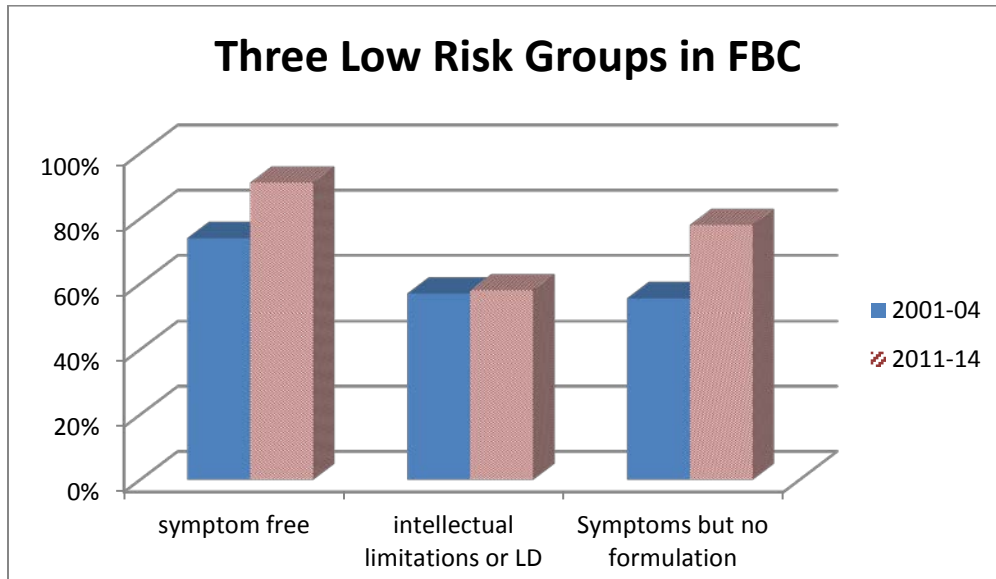
The category, symptom free, refers to children who have recovered and are no longer displaying the symptoms presented on admission. The symptom free category also includes infants and other young children who are not presenting any symptoms currently. The proportion of symptom free children placed in FBC has increased across the decade (74% to 91%). The proportion of children with intellectual limitations or diagnosed learning disability but no other comorbid conditions placed in FBC has not changed (57%). The proportion of children placed in FBC who are presenting symptoms but with no clinical formulation has increased (56% to 78%).

Placement trends are leading to an older, more disturbed child in Group and a younger less risky population in FBC.

Age when First Placed by Type of Care

Age when first placed
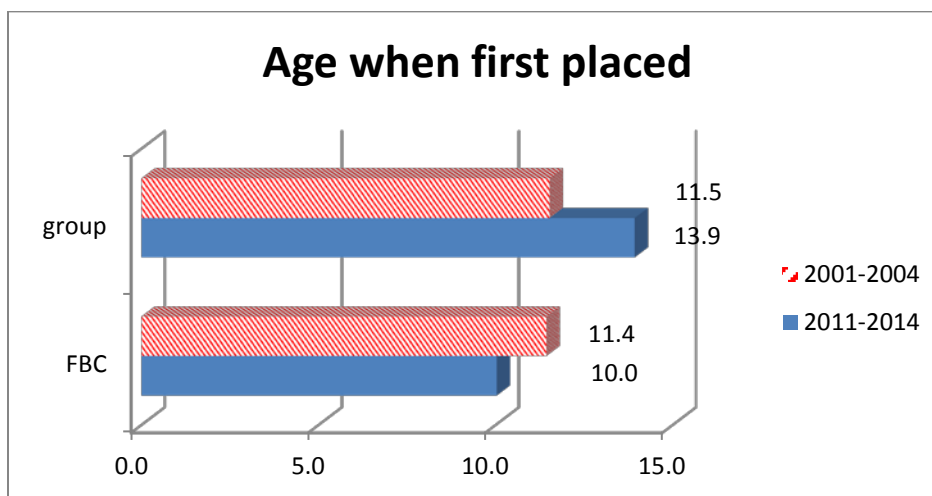
The children in Cohort 2 that are being placed in Group are 3.9 years older than the children being placed in FBC (**Group** m = 13.9 years, **FBC** m = 10.0 years, diff = 3.89, t = 13.238, n = 814, p = .000).

Findings in Relation to Positives

(1) The children in Cohort 2 are becoming more resilient compared to a decade earlier as evidenced by higher scores on tests measuring positives (i.e.) attachment. The improvements on positives (53%) are occurring in the first year of service to a large degree and continue to grow at a slower pace every nine months until year three.

Findings in Relation to the Use of Family Based Care

(1) Across the decade, the use of FBC has increased by 20%.

(2) Children in FBC are 3.9 years younger than children in Group. The average age on admission in FBC is ten years compared to 13.9 years in Group.

(3) The percentage of high risk groups (psychiatric disorders, autism) has decreased in FBC and is now found in higher frequencies in Group.

(4) The percentage of children who are symptom free and children with symptoms but who are still in need of comprehensive assessment have decreased in Group and increased in FBC.

(5) The response of individual children to positives varied by agency not by the type of placement within the agency (FBC or Group).

(6) The percentage improvement in positives varies according to how the comparison is set up (i.e.) younger children improve to a larger degree compared to older children.

(7) Children are making progress in all sub-groups.

3.3   Functioning

OARTY members use eight different measures of functioning to assist them in providing evidence based practice. Combining the effect sizes for all measures, we found an improvement of 22% in the first year of service. We found no evidence of a different response to treatment by service delivery variables, type of care and agency, and few differences by type of child. The response by individual children in terms of functioning follows a different path compared to the response on positives.

The composite variable, functioning, was computed from the Children's Global Assessment Scale, the Level of Care, academic performance, the adaptive functioning scale of the Adaptive Behavior Scale, the peer problems scale of the Strengths and Difficulties Questionnaire, the Piers-Harris Self Concept Scale, the child self-report from the Behavior Assessment System for Children (BASC) and the parent scales of the BASC. Two instruments, CGAS and Level of Care, were used for multivariate analysis.

CGAS: Functioning in the Major Roles of Life

Role adaptation across time
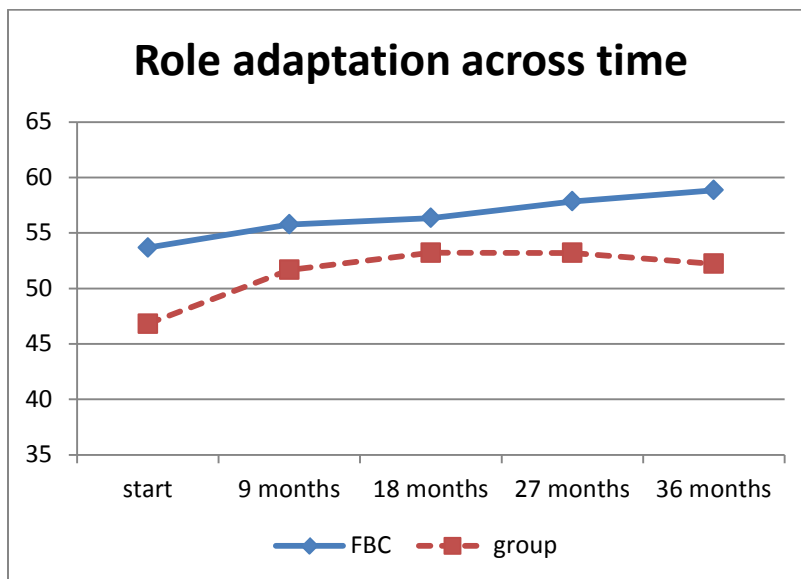
There is a statistically significant difference ($p = 00$) in each wave of testing for children in FBC and Group. Children in Group make more progress from start to finish (difference = 5.4) compared to children in FBC (difference = 5.2). Children in FBC begin their treatment at a higher level of functioning (**FBC** m = 53.7, **Group** m = 46.8, diff = 6.9, t = 8.784, n = 2,570, p = .000). Both groups improve in social role functioning.

Role adaptation across time
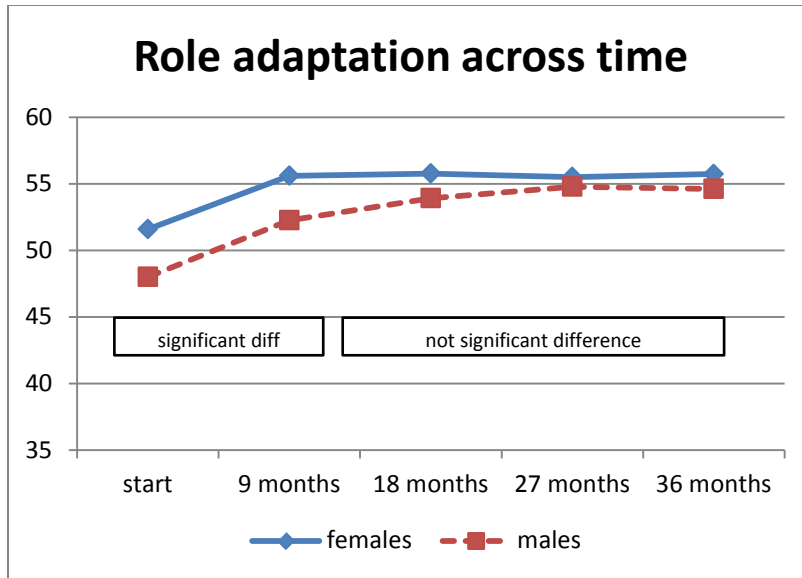
Males start their treatment with significantly lower functioning in the roles of student, family member, community member compared to females. (**Females** m = 52, **Males** m = 48, diff = 3.6, t = 4.256, n = 2,733, p = .000). In the first year of service, males catch up with females and both genders level off when CGAS = 55.

Level of Care: Dependence on Adults for Tasks of Daily Living

All groups of children improve their dependence on adults to perform the tasks of daily living by 28% across three years of service. There is no difference in the outcomes of children in FBC or Group. There is a large difference in the case of children with complex neurodevelopmental disorders.

Children with or without complex Neurodevelopmental Disorder
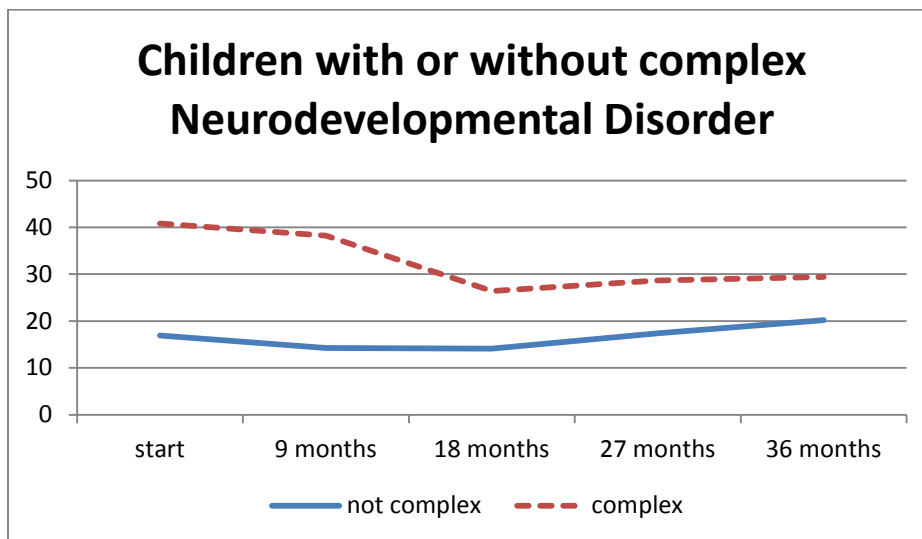
Children with complex neurodevelopmental disorders, (i.e.) FASD, autism, medically fragile and dual diagnosis, start their treatment very dependent on adults to perform 22 tasks of daily living from getting dressed to crossing the street. Children with complex NDD have an

average score of 41. After three years the children had a score of 29, which represents a 23% improvement. Although this number is lower than the effect size for all children (28%), statistical tests found the difference was not significant.

Aboriginal Children compared to Mainstream cultural groups
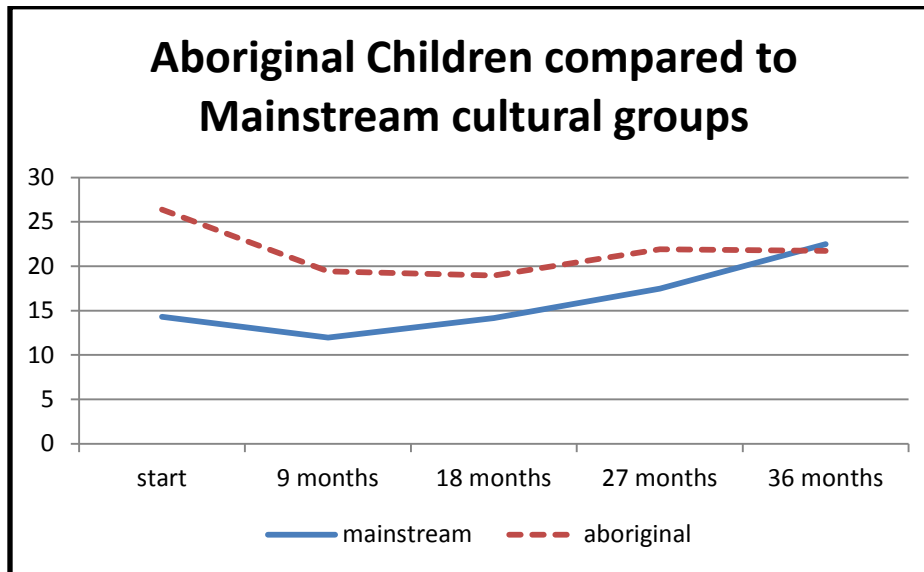
Aboriginal children display a significant difference compared to children from mainstream cultural groups when they start their treatment. The effect size at nine months is significant. (**Aboriginal** m = 32%, **Mainstream** m = 8%, diff = 24%, t = 2.572, n = 448, p = .012)

The age at the time of placement affects how much the child improves in functioning in year one. (F-ratio = 5.313, n = 1,021, p = .021; standardized co-efficient = .072, t = 2.305, p = .021)

The full group of children and all sub-groups make significant improvement during all waves of testing on level of care instrument. Various sub-groups start at different levels, but the percentage improvement and the direction of change is quite consistent throughout.

Findings in Relation to Functioning

(1) Children make an improvement of 22% in functioning based on instruments used in this study

(2) There is no evidence of a different response to treatment by type of care, agency and type of child.

(3) Different sub-groups start within a wide range on their level of functioning but most groups to end up with a similar effect size after three years of service.

(4) Children who were older when placed make more improvements in functioning than children who are younger when placed. The first test of functioning is no different statistically by age. This is the first occasion in our research when older children appear to respond better than younger children.

4.0 Risk

Five tests were used in this study to assess risk. They are the Connor's Global Index (CGI), the Symptom Agreement Scale (SA-45), the Feelings, Attitudes and Behavior scale for Children (FAB-C), the Strengths and Difficulties Questionnaire (SDQ) and the degree of adversity and trauma checklist.

Relationships between Risk and Resilience

(1) There is a significant relationship between the composite risk score and composite functioning score (F-ratio = 30.243, p = .000, standardized regression co-efficient = .175, t-score = 5.499, p = .000)

   a. This means that treatment effects in the first year of service for measures of risk and functioning are directly correlated. Outcomes in either domain predict outcomes in the other.

   b. When the composite effect size for functioning is regressed with the effect size for the Connors' Global Index with children at-risk, the relationship is stronger (F-ratio = 44.292, p = .000, standardized regression co-efficient = .306, t-score = 6.655, p = .000)

(2) There is no relationship between the composite score for positives and risk.

   a. This means that risk and positives are independent. A child can make progress in one but this does not predict what their outcomes will be on the other.

   b. Because they are independent, positives are more powerful agents of resilience. Therefore children with serious intractable impairments in terms of ADHD, psychiatric illness, brain damage can still acquire positives and can still build resilience with which to adapt better to the disorder.

Positives, i.e. attachment, close friends and happiness, can develop regardless of the number of risk factors, (i.e.) abuse, parental criminality, parental mental illness, the child may have experienced. Positives can develop in children with high or low scores on risk screening tools. Positives can develop in children regardless of diagnosis.

The definition of resilience is that the individual who is strong on positives can bounce back from adversity, trauma and mental illness. Individuals who are strong on positives cope better even if they cannot make the adverse situation (i.e.) earthquake that destroys your home, go away. Individuals with mental illness have a better prognosis, even if the illness is life-long and severe.

The reason why positives improve prognosis and our ability to bounce back is that positives change the inner working model of our mind. People who are strong on positives think about the world differently than those who lack positives.

---

(3) There are few relationships between positives and functioning. Children whose functioning is very low due to organic or environmental factors can still become more resilient through growth in positives. The exception to this rule is confined to positives and academic performance, a component of the functioning scale

   a. The effect size of the caring scale of the PBI, is significantly related to the academic performance in math (F-ratio = 4.887, p = .031)

   b. The effect size of the positives scale is significantly related to academic performance in reading (F-ratio = 6.709, p = .013)

   c. There is some interaction between the cohort and positives. Together they are related to academic performance in reading (F-ratio = 3.642, p = .033)

When a child acquires more positives, he/she does better in math and reading. This may happen because positives changes the way a person thinks; it changes the person's inner working model of the world. Academic performance is enhanced by better thinking.

The interaction between cohort, positives and performance in reading is harder to explain. We have shown that children are developing more positives in recent times compared to the children placed in the 2001-2004 period. The interaction effect with reading shows that this changing pattern of positives has made a difference in reading performance.

Functioning is defined as the ability of the individual to cope with the demands of social life, in the family, school and community as indicated by better marks in school and better scores on instruments measuring functioning. Positives and functioning are not correlated which means these two constructs are separate entities.

Treatment gains in functioning predict lower risk scores. This relationship occurs simultaneously so the cause and effect relationships may be bi-directional.

This finding may be explained as follows: Some children learn to function better in the home, school and community and get better marks in school. As a result, they start to feel better. Their emotional symptoms decrease and their risk scores decline. It is also possible some children start to feel better first and think more clearly about their situation. They may start to pay more attention in school. As a result, they start to function better.

The research does not provide enough data to validate the explanation offered above.

The impact of functioning on risk and impairment is discussed in section 5.0. Resilience is a combination of positives and functioning.

4.1   Degree of Adversity and Trauma

adversity	2001-2004	2011-2014	significance
deprivation poverty	31%	33%	ns
sexual abuse	25%	13%	****
physical abuse	41%	27%	****
suicide in family	5%	4%	ns
family member in jail	25%	17%	****
family member in psych hospital	15%	9%	****
family member has intellectual deficit	14%	9%	****
family member substance abuse	46%	38%	****
family member was raped	13%	6%	****
child with substance abuse	14%	15%	ns
violent family	9%	5%	****
sex assault in family	3%	2%	ns
child acquired brain injury	10%	7%	****
long term school failure	44%	33%	****
parent was teenage mother	20%	10%	****

A significant shift has occurred in eleven risk factors of adversity and trauma in the population of children served in 2011-2014 compared to a decade earlier. Some of these changes are clearly linked to changes in the quality of life of Ontario as a whole, i.e. the decrease in teen mothers.

The average number of risk factors has decreased significantly. (**Cohort 1** m = 3.2, **Cohort 2** m = 2.3, , diff = .89, t = 7.453, n = 1,799, p = .000)

Average Number of Risk Factors by Type of Care
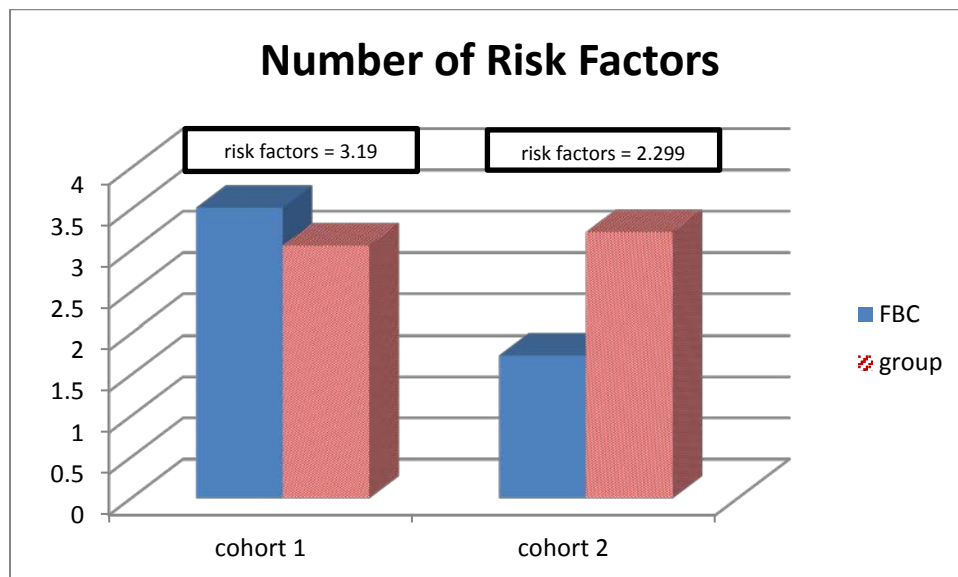
The number of risk factors carried by children placed in Group has increased by 5% and decreased by 51% for children placed in FBC across the decade. This is another finding which suggests that the population of children placed in FBC has less risk compared to children in Group. At the same time, the percentage of children placed in FBC has increased by 20% with a corresponding decrease of children placed in Group.

The only way these two facts can co-exist is that the entire population of children being placed has fewer risk factors in recent times compared to the years 2001-2004. This observation runs contrary to the impressions of people serving children in Group and FBC.

4.2 Connors' Global Index

The Connors' Global Index (CGI) is a measure of impulsiveness, hyperactivity and inattention. High scores on the CGI are indicative of ADHD or psychological distress. The content is a robust indicator of risk [5], [6].

Profile on Admission

The findings are:

(1) The recent cohort displays significantly higher scores on the CGI compared to a decade earlier. (**recent** m = 70, **decade** m = 67, diff = 3.348, t = 2.712, n = 575, p = .007). This test suggests that the recent group of children placed are more at-risk of long term mental illness compared to a decade earlier.

(2) Children placed in FBC are more at-risk compared to children placed in Group. (**FBC** m = 71, **Group** m = 69, diff = 1.811, t = 2.264, n = 1,525, p = .024)

(3) Children with complex needs are more at-risk compared to children who are not. (**yes** m = 73, **no** m = 70, diff = 3.458, t = 3.12, n = 1,380, p = .002)

(4) Females are more at-risk than males. (**females** m = 73, **males** m = 69, diff = 4.055, t = 4.888, n = 1,566, p = .002)

In summary, children at-risk of adverse lifespan outcomes on the Connors' Global Index are more likely to be found in the recent cohort, to be female and with complex neuro-developmental disabilities (NDD). The children at-risk are more likely to be found in FBC.

The term *children at-risk* is operationally defined as the probability that the child at-risk will be diagnosed with a mental disorder. There are many factors that place a child at-risk in this sense. The factors that increase the risk include hyperactivity, impulsiveness and inattention, measured by the Connors' Global Index. There are other factors that increase risk. These are measured by other instruments used by OARTY members.

---

[5] Rutter, Michael and Seija Sandberg (1985), "Epidemiology of Child psychiatric Disorder: methodological Issues and Some Substantive Findings", *Child Psychiatry and Human Development*, 15(4), 209-233

[6] Rutter, Michael (1995), "Developmental Psychopathology: concepts and prospects", in *Developmental Psychopathology*, Dante Cicchetti & Donald Cohen (eds), J. Wiley, New York, 209-237

Response to Treatment

Changes in ADHD symptoms
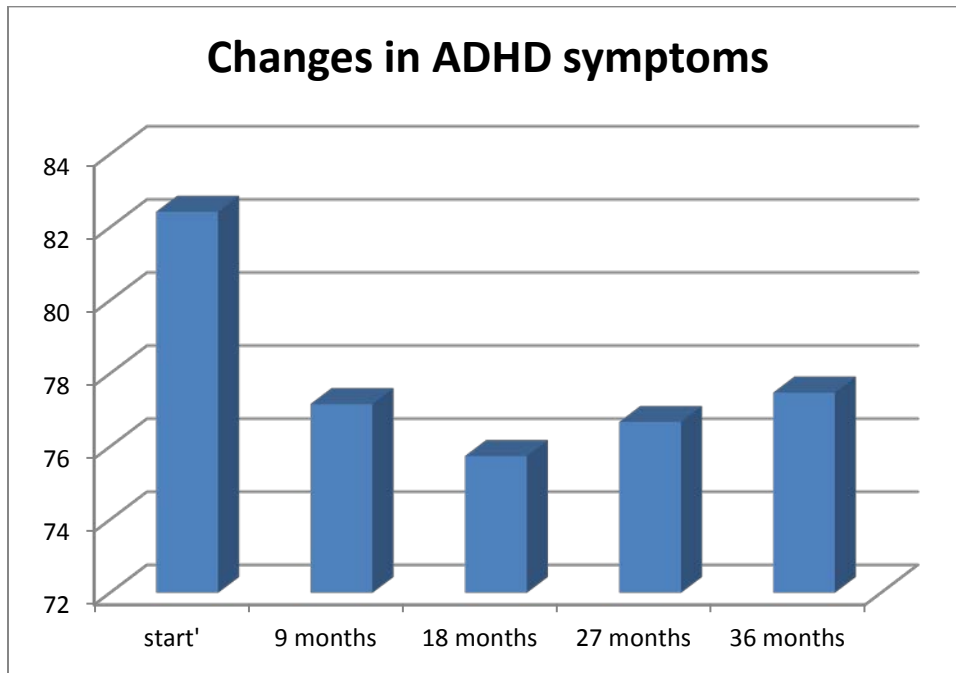
Children who were in the clinical range improved by 55% in the first year of service. The vertical axis represents t-scores on the Connors Global Index. Children continued to improve by another 12% at 18 months, but fell back. At the end of 36 months children had improved in total by 44% compared to when they first tested. There was no difference in the outcomes by:

(1) type of placement: FBC or Group
(2) year when placed, i.e. 2001-2004 compared to 2011-2014
(3) gender: except that females tended to give back their gains at 27 months (-53%)
(4) aboriginal children compared to children from mainstream cultures
(5) children with a history of prior placements
(6) children with or without psychiatric diagnosis

Two factors predict improvements on the CGI: children with or without complex neurodevelopmental disorders. (**complex NDD** m = 19%, **not complex NDD** m = 59%, diff = 41%, t = 2.437, n = 479, p = .017) and the effect sizes on functioning (F-ratio = 44.292, p = .000).

There is no correlation between the attachment scores, positives in general and degree of impairment with changes in the CGI over time. Changes on the CGI cannot be predicted from qualities in the service that are tracked by OARTY (i.e.) type of placement, cost, decade when served. There is a huge correlation between the effect sizes for functioning and the CGI. This shows children must lower symptoms of hyperactivity and inattention in order to improve in their social competence and in academic performance..

Explanation of Non-findings on ADHD Symptoms

Since high scores on the CGI is the most common risk factor affecting 52% of children, we expected to see many differences in the treatment effect between groups. We found very few differences between groups in the treatment response for the CGI.

All children placed with OARTY agencies are provided with medical consultation and would likely have a trial on medication if the consulting physician recommends this. Most of the non-findings above could be explained by the power of medication. Medication for ADHD works equally well in different agencies, in FBC or Group and in different age groups of children. Children with complex NDD (autism, FASD, dual diagnosis) do not improve as much as those who do not have these conditions. This is a reasonable given their brain damage. The different path that females take on ADHD symptoms has no explanation.

4.3   SA-45

The SA-45 is a measure of psychiatric symptoms endorsed by the young person. Often the first test is more than one year after admission, since the young person must be a teenager before taking the test. The instrument is a gold standard for population-based epidemiological surveys of the presence of mental illness. The data from this instrument predicts which children will develop a diagnosable illness and it also measures changes if the child has already been diagnosed. High scores are interpreted as an indicator of "caseness" or people similar to the profile of clients in a mental health service. The findings are:

Profile on the 1st Administration of the SA-45

(1) Cohort 1 has a higher score than the more recent cohort. (**Cohort 1** m = 57, **Cohort 2** m = 52, diff = 4.876, t = 3.828, n = 326, p = .000)

    a. Ten years earlier, 22% of children tested were above the threshold of clinical significance; in more recent years, 10% of children tested initially were above the clinical threshold.

(2) There is no difference by gender using the SA-45.

(3) Twenty six percent (26%) of children in Group are above the clinical threshold for at-risk children compared to 20% of children in FBC

(4) Children in FBC have a higher mean score on the SA-45, but the difference is on the borderline of clinical significance (**FBC** m = 73.1%, **Group** m = 70.8, t = 1.909, p = .06)

In summary, the child at-risk of mental illness as measured by the SA-45 is more likely to be placed in Group. Children with ADHD as measured by the Connors' Global Index are more likely to be in FBC.

The SA-45 is an adolescent self-report measuring feelings of depression, anxiety, anger and thoughts about persecution. These feelings are typical of the emotional problems of children placed in Group.

Response to Treatment

Psychiatric Symptoms (SA-45)
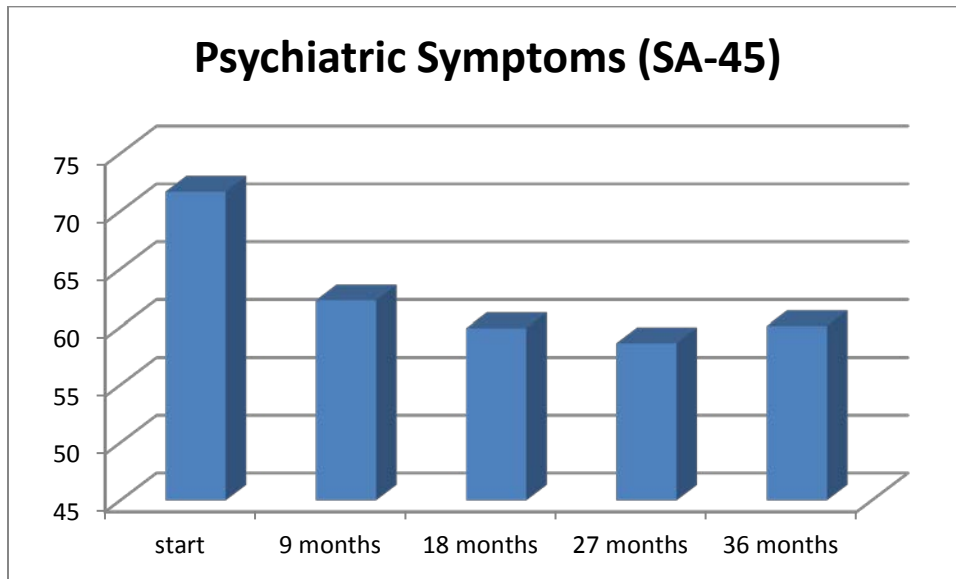
The at-risk children on the SA-45 start in the clinical range and drop to borderline range nine months after the test is administered, a 114% improvement. By 18 months, these at-risk children are in the normal range with further results indicating no clinical significance.

(1) There is no difference in outcomes by gender

(2) There is no difference in outcomes comparing children from Cohort 1 with Cohort 2

(3) Children in FBC improved to a greater degree in the first year of service by a larger margin than children in Group (**FBC** m = 160%, **Group** m = 85%, diff = 74%, t = 2.169, n = 126, p = .034)

    a. A higher percentage of children in Group (26%) were above the clinical threshold on the 1st test, but the average score of the at-risk population in Group (70.8) was lower than the children in FBC (73.1)

    b. Children in both Group and FBC had the same score at 9 months.

4.4   FAB-C

The FAB-C is a measure of antisocial attitudes, negative feelings, poor peer relationships and misbehavior. It is a self-report 48 item questionnaire designed for school age children. T-score above 65 on the Total Problem Index suggest poor affect and behavior regulation.

Profile on Admission

(1) There is no difference in the FAB-C Total Index comparing Cohort 1 and Cohort 2.

(2) Children with a history of prior placements do not have higher FAB-C scores

(3) Aboriginal children have the same FAB-C scores as children from the mainstream cultures

(4) Children with complex needs have the same FAB-C scores as children who do not have lifelong complex neuro-developmental disorders.

(4) Children placed in Group have higher FAB-C scores than children placed in FBC (**Group** m = 59, **FBC** m = 56, diff = 3.351, t = 1.625, n = 330, p = .080)

(5) There is no difference by gender.

High scores on the FAB-C are a risk factor for future mental illness. The FAB-C results parallel the findings of the SA-45 which indicate that children in Group are more at-risk of future mental illness than children in FBC. Children in Group have less affect and behavioral regulation compared to FBC.

Response to Treatment
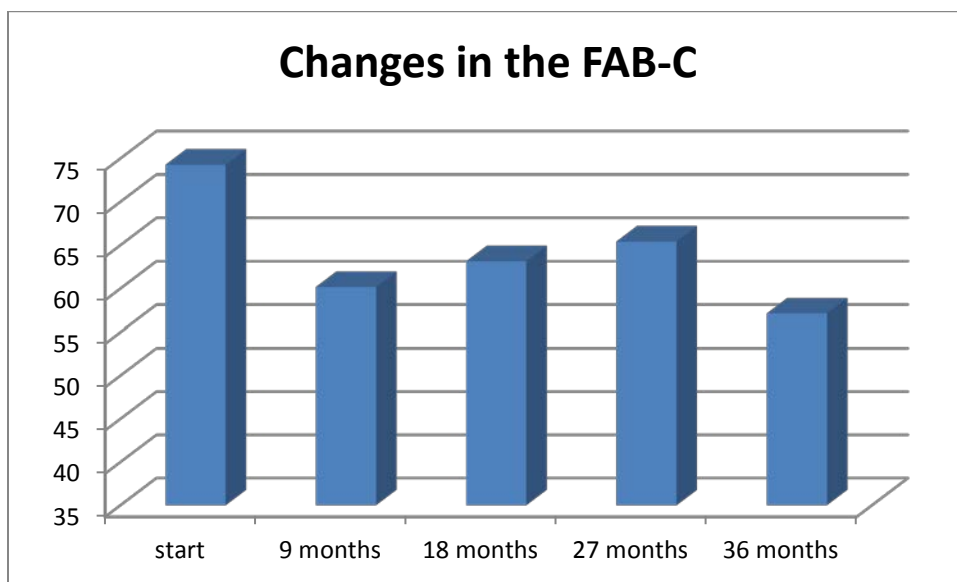
The children at-risk of mental illness as identified by the FAB-C start in the clinical range and improve to normal range nine months after the test is administered, a 162% improvement. Their symptoms increase again and decrease to their lowest point at 36 months. Across three years, these at-risk children make a 210% improvement.

(1) The type of placement made no difference on the amount of improvement. The difference is a higher percentage of children Group were in the clinical range compared to children in FBC on admission.

(2) Cohorts 1 and 2 made no difference on the amount of improvement

(3) Gender made no difference on the amount of improvement

(4) Aboriginal vs mainstream made no difference on the amount of improvement

(5) History of prior placements made no difference on the amount of improvement

(6) FASD makes a large difference (**FASD** m = 36%, **Not FASD** m = 173%, diff = 136%, t = 3.266, n = 61, p = .015)

a. This finding is consistent with the literature on FASD about their inability to regulate their feelings or their behavior

(7) Children with or without complex NDD make no difference on the amount of improvement

(8) Children with or without a psychiatric diagnosis make no difference on the amount of improvement

(9) The number of distinct adverse life events or trauma make no difference on the amount of improvement

(10) Other test scores, including positives, functioning made no difference on the amount of improvement

Combining the SA-45 and the FAB-C

The SA-45 and the FAB-C are used alternatively for different age groups. The percentages of children above the clinical threshold for each instrument can be added to provide an estimate of the percentage of children with poor affect and behavioral regulation and psychiatric symptoms.

The percentage of children above the clinical threshold is 50% in Group and 44% in FBC. Children in both types of care make a very large improvement on the FAB-C (162%) and the SA-45 (**FBC** m = 160%, **Group** m = 85%).

Summary on the Nature and Measurement of Risk

A family background marked by adversity and trauma is a reliable predictor of future mental illness especially if the child has four or more different risk factors. In addition if the child has certain behavioral and emotional symptoms, there is a probability that he will eventually be diagnosed with a psychiatric disorder. The more symptoms the child has, the higher the probability. OARTY members use four risk screening tests (i.e.) SDQ, CGI, FAB-C and SA-45 to scan for the following symptoms. Please see the appendix for literature supporting the selection of instruments.

- hyperactive and impulsive
- poor affect regulation
  - anxiety
  - depression
  - hostility
  - self hatred
- poor behavioral regulation
  - antisocial attitudes
  - lying
  - poor peer relationships
  - rule breaking behavior
  - thoughts of persecution

Fifty-one percent (51%) of children in the OARTY data repository have been administered a risk screening measure.

- 58% were found to be at-risk of mental illness the first time they were administered.
  - 19% of the children at-risk were no longer at risk one year later

Twenty percent of the children at-risk are in the normal range within one year. The remaining group of children at-risk improve significantly in subsequent years.

Findings in Relation to Positives and Functioning

(1) The children in this study made rapid and large improvements in resilience. Resilience is indicated by positives and functioning.

(2) The children continue to improve on measures of resilience over the three years of service covered by this report.

(3) The individual agency and clinical profile of the child are the primary factors which predict differential outcomes on measures of attachment. This finding was discovered through multivariate analysis. In a single variable analysis of variance, children in FBC improved by 80% compared to 36% for children placed in Group. This illustrates that children in FBC are improving to a very large degree but it would be incorrect to assume that the model of care was causing the improvement. Correlation is not evidence of cause and effect. Qualities of the child contribute to his /her response to treatment and the content and process of treatment itself also make a difference.

(4) There is no evidence of a different response to treatment on measures of functioning by service delivery variables, type of care and the individual agency. There are few differences by type of child.

  a. This finding suggests that children improve in terms of social competence, independence and academic performance as a result of the child's social context not because the type of placement.

  b. The good news for resilience is that children improve in functioning independent of their clinical condition. As a result, better functioning is able to buffer children who have a chronic clinical condition.

(5) Different clinical sub-groups start far apart on their level of functioning but all groups appear to end up with a similar effect size after three years of service.
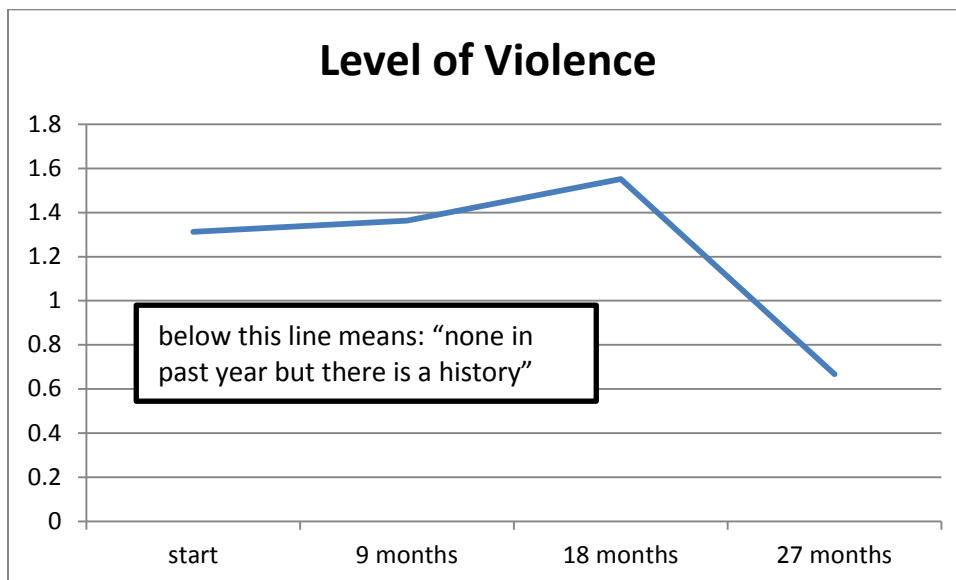
Findings in Relation to the Risk

(6) One half (51%) of children in the OARTY data repository have received a risk screening assessment and 58% were found to be at risk of mental illness.

(7) The children made large to very large improvements in the 1st year of service on all of the risk screening measures. Children consistently improve in terms of risk for mental illness to a larger degree than they improve in positives or functioning. Children placed are far less symptomatic after a year of treatment compared to their status on referral.

(8) Children in FBC show more symptoms of ADHD on their first test of the Connors' Global Index. They made a 55% improvement in the first year.

(9) A higher percentage of (26%) of children in Group are above the clinical threshold on the SA-45 indicating lower affect regulation and lower behavior regulation on their first test of the SA-45. They improved by 85% after one year in Group.

(10)   Twenty-four (24%) of children tested on the FAB-C were above the clinical threshold indicating poor affect and behavioral regulation.

(11)   Children at-risk made a very large improvement on the FAB-C in the first year (162%). The response to treatment did not vary by FBC or Group.

(12)   Combining all measures of risk, 1/5th of high risk children are in the normal range after one year of service and continue to improve thereafter.

---

(13)   There is a significant relationship between the composite risk score and composite functioning score

    a. As children improve in terms of social interaction, dependence on adults and academic performance, they present fewer symptoms that indicate risk for mental illness.

    b. There is a two-way interaction between risk and function. Children learn to function better when they have better affect regulation, better behavior and are more attentive. Therefore, risk and function affect each other.

(14)   Functioning is a dimensional concept. Children can become better socially and in school without limit.

(15)   There is less dimensionality in risk. Once the child's scores are below the threshold indicating risk, changes in the scores are meaningless.

5.0 Impairment

Twelve instruments are designed for children with diagnosed disorders and/or dangerous or self-harming behavior. Once the child has been diagnosed and/or is already exhibiting dangerous behavior, risk screening is less relevant. The composite score for *impairment* is a combination of the internalizing and externalizing scales of the parent rated Behavior Assessment System for Children (BASC), the behavior reduction scale of the Adaptive Behavior Scale and serious behavior problems checklist of the Inter-Rai instruments, used in Ontario Psychiatric facilities.

The impairment composite score shows no treatment effect in the first year of service. In several component scales, children tended to get worse in the 18 months and then improve by the 27th month. This is illustrated by the level of violence scale.

Level of Violence

below this line means: "none in past year but there is a history"

start   9 months   18 months   27 months

(1) There is a significant relationship between type of placement and the composite impairment treatment effect size: (**FBC** = -16%, **Group** = 27%, diff = 43%, n = 102, t = 2.532, p = .015)

   a. Children in FBC appear to get worse to a small degree in the 1st year of service while children in Group make a moderate improvement.

(2) The composite impairment score is based on 155 cases tested at least twice. This is a small sample size compared to the sample sizes for the positives, functioning and risk.

   a. This means that we cannot test for any relationships between resilience and impairment

6.0   Resilience and Risk: Improvements with Time

OARTY members have successfully implemented a system of outcome measurement. The members have followed 2,493 children through years of service with multiple tests on the same set of instruments. Approximately 25% of all children currently in service are being followed by their agency.

After years of data collection, OARTY is now in a position to calculate the percentage of improvement or decline in outcomes across four critical domains: positives, functioning, risk and impairment. The domains that OARTY members measure fit perfectly into the resilience model. The data shows that

1. Children are acquiring greater resilience while in care

2. The nature of resilience includes positives such as attachment to parents and caregivers, positive peer relationships, happiness and client satisfaction. These qualities are improving at the rate of 53% in the 1st year of service and continue thereafter.

3. Our operational definition of resilience includes functioning such as social competence, independence and better marks in school. These qualities are improving at a rate of 22% in the 1st year of service and continue thereafter.

4. The capacity of children to develop resilience is independent of diagnosis, developmental challenges, adversity and type of care. Children in Group and FBC respond equally well.

5. A large percentage of children (58%) are at-risk of mental illness as measured by evidence based risk screening tools. These children are becoming less at-risk to a very large degree ranging from 55% to 160% in the 1st year of service.

6. The treatment effect for functioning and risk are highly correlated. The two domains support each other in a common cause for a better future for the child. This is predicted by the resilience-risk model.

OARTY members care for children with severe impairment in behavior, developmental conditions and psychiatric disturbance. Children who present severe impairment are beyond a question of risk. Reducing impairment for severe cases takes much longer than improvements in resilience and risk. The data shows that indicators of impairment typically get worse for the first two years of treatment after which the typical child in our care starts to respond.

In future research studies, OARTY plans to examine the relationships between resilience to the reduction of impairments.

Robert Fulton, M.S.W., Research Consultant

Appendix A: technical details

1.0   The State of the Data Repository

The Ontario Association of Residences Treating Youth (OARTY) maintains a data repository of anonymized clinical information on the clients served by member agencies. The client data comprises of:

- Clinical profiles drawn from random samples of all clients in service from 98 agencies
- Clinical profiles of all clients enrolled in service for ten agencies
- The results of a program of testing children every nine months from the date of admission until discharge
    - Initially, agencies were testing clients with eight measures: Degrees of Adversity and Trauma, CGAS, Connors Global Index, SA-45, FAB-C, Level of Care, Daily Stressors and the Parental Bonding Instrument
    - Currently, agencies use combinations from a basket of 49 instruments
- business and financial data at the level of
    - Agencies (n = 213)
    - Programs (n = 476)
    - group homes (n = 261)
    - foster parents and staff.

The repository contains clinical profiles on 4,616 unique clients from 98 agencies. A total of 1,361 new clients have been added this year for PIC 6.

1.0   Research Design: 2015

This study will focus on children admitted to the member agency during the years 2011 to 2014. We will report on the child's clinical profile and his/her scores on outcome measures. The outcome measures address into four content areas: positives, functioning, risk and impairment. The instruments associated with each content area are discussed in section 3.1.

For every child tested, the instruments are organized in date order. The resulting test sequence is labelled, wave 1, wave 2, wave 3 and so on. The difference in the scores for each pair of tests in the sequence is calculated. As a result each child is his own control. We are comparing each child's response to treatment in relation to the baseline or 1st test for that child.

The clinical profile of the children in the period 2011 to 2014 is compared to an earlier period of children placed from 2001 to 2004. No child appears in both groups; the list is completely exclusive. For clarity the earlier comparison group is called cohort 1. The primary focus of the research is referred to as cohort 2.

1.1 Cohort 1: (placed in 2001, 2002, 2003 or 2004)

We have multi-wave test data on 761 children from 52 agencies

1. A Circle of Support
2. Albert Powell Home
3. Amity Residential Treatment
4. Arden Court Children's Residences
5. Ariss Place Residential Care & Treatment
6. Avalon
7. Bairns Wee Croft
8. Bayfield Treatment Centres
9. Bertzel Manor Inc
10. Bluewater Family Support Services Inc.
11. Bob Rumball Associations for the Deaf
12. Care Management Group Home
13. Carpe Diem Foster Homes
14. Cavan Group Home
15. Community Support Housing
16. Connor Homes
17. Cornerstone Group Homes
18. Cricket Hollow Inc
19. Delta Lee Homes Inc
20. DFK Residential Services
21. DNAB Residence
22. Doris's Nest
23. Eagle's Nest: A Place to Soar Inc.
24. East York Residential Services
25. Enterphase Child & Family Services
26. Elk's Boys Home
27. Falconhurst
28. Family Tree Youth Services
29. Fernie House Child & Youth Services
30. Flourish Children's Home
31. Grand Riverview
32. Guardian Youth Services
33. Hawk Residential Care & Treatment Homes/Hollyhawk Treatment Foster Care Services
34. Kennedy House Youth Services
35. Little House Residential Care Services Inc.
36. Milestone Foster Homes
37. Mon Ami Children's Services
38. Mutual Support Systems
39. Nairn Family Homes

40. Oakdale Child & Family Services
41. Phoenix Residential Treatment Services
42. Pioneer Youth Services – K/W
43. Partners in Parenting
44. Quinte Children's Homes
45. Ridgeview Children's Home Inc.
46. Sinclair Children's Residence
47. J & D Residential Services/S.T.A.R.T.
48. Stewart Homes
49. Storey Homes
50. Summit Human Services
51. Susie's Place Ltd.
52. Woodcock Youth Centre

1.2 Cohort 2: (placed in 2011, 2012, 2013, or 2014)

We have multi-wave data on 1,070 children from 23 agencies.

1. A Circle of Support
2. Ariss Place Residential Care & Treatment
3. Bayfield Treatment Centres
4. Care Family & Children Services
5. Connor Homes
6. DLC Residential Services
7. Eagle's Nest: A Place to Soar Inc.
8. Enterphase Child & Family Services
9. Hawk Residential Care & Treatment Homes/Hollyhawk Treatment Foster Care Services
10. Jen's Place
11. Little House Residential Care Services Inc.
12. Mon Ami Children's Services
13. Morningstar Family Ministries of Canada Inc.
14. Mutual Support Systems
15. Nairn Family Homes
16. Phoenix Residential Treatment Services
17. Partners in Parenting
18. Quinte Children's Homes
19. Sinclair Children's Residence
20. Stewart Homes
21. Summit Human Services
22. Terrace Youth Residential Services
23. Youth Connections

1.4   Total Study Population

There are 1,831 unique children in the comparison group and the study group from the years 2012 to 2014. There are other clients served in years not included in the two cohorts. The data repository includes 2,493 unique children with multi-wave test data from 98 agencies. Ten agencies have enrolled all of their clients in the study as of 2015. Eighty-eight agencies have enrolled a sample of the clients in the longitudinal study.

1.5   Larger Context

The proportion of children covered by the data is indicated by the system wide statistics for 2013. In that year, 64 member agencies served 2,791 children. We collected outcomes from 684 of the children served in 2013, representing 25% of the total.

2.0   Previous Findings from the OARTY Longitudinal Data

OARTY has been supporting a longitudinal outcome study of children and youth receiving services since 1993. This study began with five agencies initially and has since expanded.

In 2012, OARTY published the results of the first fifteen years of data on the original longitudinal design with five participating agencies. In order to compute the outcomes, the response for each client was examined in relation to each outcome measure. The difference in scores comparing wave 1 to wave 2, wave 2 to wave 3, wave 3 to wave 4 and so on was found for each measure. The difference was divided by the pooled standard deviation for the test pair. The resulting metric is a standardized change score, referred to either as Cohen's d or the effect size. International standards exist for interpreting the effect size. Depending on the value of the effect size, the treatment outcomes can be described as small, medium, large and very large.

> "For all of the measures subject to change, the data shows a strong to very strong positive treatment effect. The majority of children who started the study within the at risk zone crossed over to the range of scores indicative of a normal child."

The specific findings were:

(1) The effect sizes for the Conners' Global Index, a measure of ADHD symptoms and psychological distress were between .51, and .68. This indicates a strong positive treatment effect according to international standards for social service agencies.

(2) The SA-45 Global Severity Index is a measure of broad psychiatric symptoms. The outcomes of children demonstrated a very strong treatment effect across four years of service. During each cycle of testing, the children made clinically significant improvement in a broadly based range of psychiatric symptoms. At the end of four years, 68% of all of the children moved from the clinical range to the normal range.

(3) The FAB-C Problem Index is a measure of antisocial attitudes, negative feelings, poor peer relationships and misbehaviour. The outcomes of children demonstrated a very strong treatment effect across four years of service. The percentage of children whose scores crossed the threshold from at risk to normal scores varied from 57% to 83%; an average response rate of 69% across 3.3 years of service.

(4) The Children's Global Assessment Scale is a measure of how well children are able to function at home and school. The outcomes of children demonstrated a strong treatment effect across four years of service with the most substantial improvement occurring the first year of treatment.

(5) The Level of Assistance Scale (LOA) is a measure of how much adult support the child requires to get through the day in relation to basic tasks, such as eating, dressing and following rules. The outcomes for children on the LOA demonstrate a strong positive treatment effect, especially at the end of the first year of service. Fifty eight percent of the children in the at risk cohort ended up in the normal range. The results on the LOA indicate that the OARTY agencies are able to improve the child's independence in daily living skills to a clinically significant degree.

(6) Attachment was measured by the Parental Bonding Instrument, which has three scales; the first measuring the degree of affection the child perceives from his/her caregiver. 60% of children in the study began their placement with a very low affection score. The average for this at risk group was 30.7.

   a. The average score measuring affection for the at risk children was below the 3rd percentile of teenagers all over the world. When these children were placed in residential care, they felt that their parents disliked them intensely. The outcomes for the at-risk children demonstrated a strong treatment effect every time they were tested.

   b. The second scale of the PBI is able to identify children who feel they are neglected and unsafe when reflecting on his/her closest caregiver. 28% of the children felt they were neglected and unsafe when first admitted. The outcomes for the children who were in the at risk group demonstrated a very strong treatment effect. The largest change occurred at the end of 3.25 years of treatment. During the wave 4 to 5 test period, 94% of the at-risk cohort had crossed the threshold indicating they feel safe and were not neglected. Eighty percent of those at risk on this measure, returned to the normal range during the study.

   c. The third scale of the PBI measures the degree of unfairness, i.e., the child feels he/she was "over-controlled" resulting in the child feeling his/her parents did not understand him/her. 52% of the children felt they were treated unfairly at wave #1. The outcomes for children in this at risk group demonstrated a very strong treatment effect. The most substantial change occurred at the end of 3.25 years of treatment. During the wave 4-5 test

period, 47% of the at-risk cohort had crossed the threshold indicating they feel their needs were understood and they were fairly treated.

3.0 Administration of Outcome Tests

OARTY members have been administering risk screening instruments and outcome measures of their clients since 1996. Agencies have shared clinical profile data and test results with OARTY since then. Software, the OARTY Information System or the OIS, was developed to support the longitudinal study. Five member agencies sent regular updates to OARTY of their test data, using a computer generated unique individual identifier to keep the data anonymous. Field testing began in 1996 and the five agencies started send data on all children served between April 1998 and August 2014, when this longitudinal study came to an end. During that time, 1,061 children were administered outcome tests up to 14 times. The agencies participating were:

- Bayfield Treatment Centres
- Connor Homes
- Enterphase Child & Family Services
- Hawk Residential Care & Treatment Homes/Hollyhawk Treatment Foster Care Services
- Mutual Support Systems

Overtime the OIS was replaced by two 3rd party information systems, HOMES and Sharevision. Two web based information systems, Morningsun Software and E-CATS were developed by member agencies.

In 2012, OARTY developed new software for those members who required it. OARTY now offers its members an Excel program running data entry forms or an Access database that can be deployed in a server environment.

3.1 Consultation with Members on which outcomes matter most

During 2014, we consulted agencies representing all types of service models and client groups in order to evaluate the relevance of the outcome measures selected by OARTY. The outcomes used by OARTY cover the following domains:

1. Positives, strengths and resilience
2. Ability to function in aspects of daily living
3. Risk screening (i.e.) the probability of acquiring a diagnosable disorder
4. Degree of Impairment for children with diagnosed disorders and/or dangerous or self-harming behaviour
5. Infant Development
6. Quality of Life for the Medically Fragile
7. Quality of Family Functioning with Family Based Care settings

Agencies providing treatment foster care added the seventh domain, *Quality of Family Functioning*, as a critical area for measurement. A number of new measures were added under each domain to ensure that appropriate measures are available for all types of children and families. The measures used under each domain are evidence based tools and many are recommended by the Royal College of Psychiatry or the American Psychological Association.

(a)   Positives, strengths and resilience

Nine assessment measures are used:

- Kern's Security Scale (KSS)
  - Attachment to parent figures
  - Children 8 years to 14 years old
  - Self-report
  - norm referenced with t-scores
- Parental Bonding Instrument (PBI)
  - Attachment to parent figures
  - 13 years old to adults
  - Self-report
  - norm referenced with t-scores
- The Inventory of Parent and Peer Attachment (IPPA)
  - Attachment to peers
  - 13 years and up
  - Self-report
  - norm referenced with t-scores
- Piers-Harris Self Concept Scale
  - 9 component scales, including one that measures positive
    - Happiness scale
  - Self-report
  - norm referenced with t-scores
- Self-efficacy
  - Child's internal sense of empowerment and capacity to solve problems
  - All ages
  - Self-report
- Social Support
  - Availability of people for purposes of companionship, assistance, guidance or other types of support
  - All ages
  - Self-report
- Client Satisfaction
  - Client's perception of how helpful the resource is
  - All ages
  - Self-report
- Client Goals

- record of goals set and accomplished by the client during plans of care
- All ages
- File review
- Strengths and Difficulties Questionnaire (SDQ)
  - Six scale including one which measures positives
  - Pro-social scale

(b) Ability to function in aspects of daily living

The client's functioning was measured by nine instruments:

- Academic performance, derived from
  - IEP report cards specifying grade level equivalents in four core subjects
  - General report cards, specifying grade scores or general comments
  - Standard scores on the WIAT
- Children's Global Assessment Scale (CGAS)
  - Rating how well the client's is adapting to home, school and community
  - The score is a number from 1 to 100
  - Ten exemplars that are the standard criteria for making a judgement
  - Worker bases judgement on file review and experience with the client
- Peer Problems Scale
  - Scale #10 from the HoNOSCA, a mandatory outcome instrument by EU countries, Australia and New Zealand
  - Applies to children and adolescents
  - Worker makes a judgement of how severe the client's peer relationship problems are from no problem to severe
  - Criteria is specified and a form is provided to guide the worker to make the judgement
- Level of Care
  - Measures how much adult support is required to ensure the client is able to complete 22 daily tasks of living (e.g.) getting dressed, crossing the street
  - Adult support is measured on a five point scale from totally independent (0), daily oversight (1) .. adult does everything (4)
  - A standard component of the mandatory instrument used to assess needs for personal support workers by CCAC
- Adaptive Behaviour Scale
  - Measures how well the child is functioning in over 18 major domains
  - Covers both skills and symptoms
- Piers-Harris Self Concept Scale
  - 9 component scales, including one that measures functioning
    - I. Total self concept, which includes behaviour regulation, intellectual and school status, physical appearance and attributes, freedom from anxiety and popularity

- o Self-report
- o norm referenced with t-scores
- Strengths and Difficulties Questionnaire (SDQ)
  - o Six scale including one which measures functioning
  - o Peer problems scale
  - o norm referenced with t-scores
- Behavior Assessment System for Children (BASC)
  - o 21 scales, including five that measure functioning
    - i. Self esteem
    - ii. Self reliance
    - iii. Social problems
    - iv. Locus of control
    - v. Social adjustment
  - o Self report
  - o Norm referenced with t-scores
- BASC for parents
  - o A completely different set of items compared to the child BASC
  - o A behaviour observation rating scale completed by the parent
  - o 18 scales, including four that assess the child's functioning
    - i. Adaptive skills
    - ii. Social Skills
    - iii. Activities of Daily Living
    - iv. Communication skills

(c)   Risk screening

Six instruments measure the probability of acquiring a diagnosable disorder, either physical or mental. The instruments are:

- Degree of Adversity and Trauma
  - o Original instrument, *Sociodemographic Checklist*, was renamed
  - o List of 15 serious risk factors in the social background of the children
  - o Derived from Emily Werner's longitudinal study on the Island of Kauai
  - o The number of different types of stressors predicts future mental illness and dysfunction, rather than one particular stressor repeated any number of times[7].
- Strengths and Difficulties Questionnaire (SDQ)
  - o ages 3 years and up
  - o parent and teacher behaviour observation scales
  - o self-report form for youth 11 years to 17 years
  - o most accurate predictor of future mental illness
    - ▪ does not predict a specific type of mental illness

---

[7] Werner, E. (1989), "High Risk Children in Young Adulthood: A longitudinal study from birth to 32 years", *American Journal of Orthopsychiatry*, 59(1), 72-81

- o norm referenced with t-scores
- GAIN-SS
  - o Ages 13 up
  - o Youth Self-report
  - o Predicts substance abuse, suicidality, conduct disorder
- Behavioural Determinants of Health
  - o Current public health screening form
  - o Targets behaviours (e.g.) smoking and unsafe sex that carry a significant risk of disease in the future
  - o Worker identifies risk behaviour from observation and file review
- Everyday Stressors
  - o Semi-structured child/youth interview
  - o Interviewer discusses six domains of stress (e.g.) life change events, daily hassles, chronic stressors
  - o Score = the count of distinct stressors that are on the mind of the client at the time of the interview
  - o predicts quality of life, performance in school, attention and resilience
- Conners' Global Index
  - o The gold standard for measuring ADHD
  - o High scores also predict psychological distress
  - o Both are highly predictive of future mental illness
  - o norm referenced with t-scores

(d)   Degree of Impairment

Twelve instruments are designed for children with diagnosed disorders and/or dangerous or self-harming behaviour. Once the child has been diagnosed and/or is already exhibiting dangerous behaviour, risk screening is less relevant. The instruments measure psychiatric and emotional symptoms and the degree of dangerousness. The worker would select an instrument that matches the child's diagnosis or pattern of misbehavior.

- Behaviour Modification Data
  - o ABA is a databased intervention and continually counts the number of times that the target behaviour occurs
  - o This forms captures the ABA behavioural counts
- Serious Behaviour
  - o A measure of how frequently specific serious behaviour occurs
  - o Covers behaviour that is harmful to others, harmful to self, odd behaviour and violence
  - o Worker measures occurrence on a four-point scale ranging from never, past year, in one or two of past three days and every day for the past three days.
  - o This produces a total score of serious behaviour
- Danger to Others Checklist

- o Behaviour observation scale
- o Appropriate for foster parents or CYW
- o Checks whether or not behaviour occurred during the day across a one week period
- o Behaviours are specified from four domains: fighting, property damage and fire, hostility and provocation
- Harm to Self Checklist
  - o Behaviour observation scale
  - o Appropriate for foster parents or CYW
  - o Checks whether or not behaviour occurred during the day across a one week period
  - o Behaviours are specified from the four domains: suicide, tissue damage, risk taking and painful emotional symptoms
- Alcoholism screen
  - o Worker judgement based on criteria
  - o Five point scale: abstinent, use without impairment, abuse, dependence and dependence with institutionalization
  - o Criteria is fully defined
- Illegal Drug Use
  - o Worker judgement based on criteria
  - o Five point scale: abstinent, use without impairment, abuse, dependence and dependence with institutionalization
  - o Criteria is fully defined
- Neuro-behavioural Symptoms
  - o Behaviour observation checklist
  - o Screens for behaviour that may caused by impaired brain function
  - o Domains covered include: executive functioning, attention, communication, sensory-perception, motor skills and social functioning
  - o Generally, these behaviour are not subject to learning, but are driven by internal processes in the brain and not connected to the environment
  - o For example: "Excessive time and distress when playing a formal game with rules and other players, such as, baseball, bowling"
  - o Accommodating the child's slower cognitive processing, teaching the rules of the game and providing visual cues may eliminate this symptom without changing the underlying defects in cognition that produced the symptoms in the first place
- Feelings, attitudes and behaviour for children (FAB-C)
  - o Ages 6 to 13
  - o child self-report
  - o norm referenced with t-scores
  - o measures affect and behaviour regulation in school age children, a critical domain of resilience

- appropriate for children who do not display affect and behaviour regulation
- Symptom Agreement – 45 (SA-45)
    - Youth self-report
    - Ages 13 to adult
    - norm referenced with t-scores
    - measures the number of symptoms that the youth endorses that are mapped to DSM definitions of mental illness
- Childhood Autism Rating Scale (CARS)
    - Behaviour observation scale
    - Measures the occurrence of 15 symptoms of autism
    - The total score ranges from zero to 60
    - When the CARS total score goes down over time, this reflects a better quality of life for children diagnosed with autism
- Adaptive Behaviour Scale
    - Measures how well maladaptive behaviour
    - Behaviour observation rating scale
- BASC for parents
    - A behaviour observation rating scale completed by the parent
    - 18 scales, including three that assess the child's degree of symptoms
        i. Externalizing behaviour
        ii. Internalizing behaviour
        iii. Emotional symptoms

(e) Infant Development

Four instruments are used to assess infant development. The first three in this group were discussed above but are also appropriate for infants. Changes over time indicate better developmental outcomes for the child.

- the SDQ
    - parent behaviour observation
    - ages 3 to 4 and 5 to 10
    - norm referenced with t-scores
    - domains covered include: pro-social skills, emotional functioning, conduct problems, hyperactivity and peer problems
- Conners' Global Index
    - Ages 2 and up
    - norm referenced with t-scores
    - measures hyperactivity
    - predicts psychological distress as well impaired brain functioning
- Level of Care
    - Worker judgement of how much adult support is required to ensure that the child performs 22 tasks of daily living

- Nippissing District Developmental Screen (NDDS)
  - Different forms according to the age of the infant (e.g.) 1-2 months, 4 months, 6 months ..
  - Behaviour observation scale for parents
  - Indicators that the child is meeting developmental milestones
  - Two or more negative indicators predict the need for comprehensive infant assessment.

(f)   Quality of Life for the Medically Fragile

Clients of OARTY members who are medically fragile or physically disabled have little opportunity to improve in relation to their condition. The outcomes selected for this group measure quality of life in a variety of dimensions. Nine instruments are used with the medically fragile. One of them, the level of care, is used to assess functioning for types of cases.

- Health of Skin and Feet
  - Worker judgement based on rating scales and checklists
  - A six-point scale that measures the condition of the skin, which is at risk in this population and can be in good condition through proper care
  - A four-point scale that measures the condition of the feet which are at-risk in this population and be improved through proper care
  - A checklist of remedies and accommodations for skin and feet
- Braden Scale
  - A risk screening tool for measuring the probability of pressure sores on the skin; this provides a context for assessing the results of outcomes on skin test above.
  - The risk is expressed as a total score between 1 and 27
- Cognition and Perception
  - measures consciousness, hearing and vision, methods of communication, expressive and receptive language skills and accepting care
  - form includes both checklists and rating scales
  - a composite score can be derived
  - improvements can be observed even in severe cases
- Bladder control
  - Five-point rating scale
- Bowel control
  - Five-point rating scale
- Use of Hospitalization
  - Visits to ER
  - Admission to inpatient unit
  - Visits to day clinics
- Physician Visits
  - Days in previous 14 days when child saw a physician or was prescribed changes to medication

- Rehabilitation Service
    - Days During previous when more than 15 minutes of rehab was provided
    - 10 types of rehab are measured
- Level of Care
    - Worker judgement of how much adult support is required to ensure that the child performs 22 tasks of daily living
    - Adult support is measured on a five point scale from totally independent (0), daily oversight (1) .. adult does everything (4)

(g)   Quality of Family Functioning with Family Based Care settings

The logic of testing in this area is systems theory. Systems theory defines a foster family in the following terms:

The family is seen as an open system consisting of systems within systems (individual, marital) and relating to other systems (extended family, schools, industry and religion). There are explicit and implicit rules, plus action by members, which govern and monitor each other's behaviour. The parts of the family are inter-related, and the individual's feelings, interaction and behaviour within the family cannot be understood apart from the other members.

Agencies delivering treatment foster care provide recruitment, home studies, training, support staff and models of intervention that are anchored on systems theory. They do not emphasize treating the child directly. Rather they prefer to treat the child indirectly through the rehabilitative power of family life. Family is a powerful agent of change for better or worse in the child's life. The quality of functioning within the foster family is the active ingredient that changes the child. Extensive research family functioning has found a substantial proportion of families worldwide who are dysfunctional[8]. The research also found that the way the family system operates can be functional in one context but becomes dysfunctional in another. The best indicator of a successful family is their capacity to change and adapt.

Quality of functioning on a family level is measured by eight instruments:

- Family Assessment Device (FAD)
    - Each parent rates their family from their personal vantage point
    - The FAD is made up of 60 questions rated on a four-point scale
    - Measures problem solving, communication, role performance, affective responsiveness, affective involvement (i.e.) boundary problems, behaviour management skills and general functioning
- General Functioning scale
    - A component of the FAD
    - May stand alone as an outcomes
    - norm referenced with t-scores
- F-COPES

---

[8] Ryan, Christine E.; Epstein, Nathan B.; Keitner, Gabor, I.; Miller, Ivan W. & Bishop, Duane S. (2005), *Evaluating and Treating Families: the McMaster Approach*, New York: Routledge

- o Each parent rates how they cope with stress in their lives
- o F-COPES is made up of 30 questions rated on a five-point scale
- o Measures their preference for five types of support plus a total support score: social support, cognitive coping (i.e.) reframing, spiritual support, mobilizing resources and passive appraisal (i.e.) avoiding the issue and recharging energy
- o norm referenced with t-scores
- Questionnaire on Resources and Stress (QRS-F)
  - o Parent rates 52 questions covering different types of stress related to a particular (i.e.) target child.
  - o Measures the impact on family life of the target child, pessimism about the child's future, child's level of dysfunction and isolation and child's physical limitations
  - o A total score can be derived for assessing the burden on the parent
- Parenting Stress Index
  - o Parent rates his/her stress levels with 25 questions, measuring
    - i. parental distress
    - ii. distress about the child's well-being and future
    - iii. parent child dysfunctional interaction
  - o norm referenced with t-scores
- Heartland Forgiveness Scale
  - o Parent answers 18 questions on a seven point scale
  - o The content of the test concerns how the foster parent perceives him/herself in the context of interpersonal conflict or mistakes
  - o A total score and three scales are produced: forgives self, forgives others, forgives situations.
  - o High scores predict people able to maintain good interpersonal relations
- Self-efficacy
  - o Foster parent's sense of empowerment and capacity to solve problems
  - o Self-report
  - o Seven questions rated on a five-point scale
  - o This version was adapted and validated by Statistics Canada for aboriginal parents
- Social Support
  - o Availability of people for purposes of companionship, assistance, guidance or other types of support
  - o Self-report
  - o Eight questions on a four-point scale
  - o This version was adapted and validated by Statistics Canada for aboriginal parents

3.2 the Number and Scope of Measures

There are forty-nine distinct instruments in the OARTY basket of outcome measures. Some of the instruments serve two or three different constructs, such as positives, functioning and risk or different client groups, such as infants and adolescents.

Different combinations of the measures above are currently used by the member agencies. Outcome evaluation is occurring primarily at the agency level. At the association level, OARTY gathers anonymized data from its members and summarizes the results.

New research methods have been developed to manage the data analysis across a wide variety of instruments. This is discussed below.

4.0 Research Methodology

Each client in the study has been tested at least twice using one or more of the instruments described above.

4.1 Set up database for longitudinal study

Two fields were added to the database for each test: wave_number and wave_type. The wave_number was populated for each client and each test result in ascending order from the date of the first test. If a client had only one record, the wave_number was left blank and wave_type was populated as "prof" or "base".

The records marked as "prof" include clients from several years ago whose scores represent a "point in time" or a profile. The records marked as "base" were completed in the past year and are still in care. As the OARTY data repository is updated, these clients will be probably be tested a second time.

For clients tested multiple times, the first test in the series has a wave-number of 1 and wave_type populated as "pre". There are no tests in the database that were completed before admission. Statistics on when the "pre" test was completed during the history of service will be reported.

The wave_type for the last test in the series for each client was populated as "post". There are no tests in the database that were completed after discharge; the "post" test means simply the last test completed.

All records were examined for missing values and impossible scores. These were deleted.

4.2 making sense of changes over time

(a) Explanation of the effect size.

The statistic used to assess changes over time is Cohen's d or the standardized treatment effect[9]. This statistic is used in meta-analysis of published research to compare the outcomes from hundreds of programs and dozens of instruments. Different tests often use different numbering systems to measure the outcome, such as a grade equivalent, a t-score, a simple count of the number of stressors, a percentage of tasks completed and so on. Cohen's d changes all the different numbering systems into a standard format on a bell curve.

Cohen's d is also known by two other words, the *effect size* and the *treatment effect*. The *effect size* is the raw difference between the mean score on admission and the mean score nine months later, divided by the pooled standard deviation of wave 1 and Wave 2. The pooled standard deviation equals the square root of ((SD Time 1 squared * (Time 1n -1) + SD Time 2 squared * (Time 2n -1)) Divided by (Time 1n + Time 2n – 2).

Cohen's d = the raw difference in scores divided by the pooled standard deviation. The formula converts the raw difference in scores into a standardized value. The guidelines for interpreting the effect size are specified in the footnote[10].

The effect size is usually expressed as a percentile. Clients who improved by one half of a standard deviation are described as improving by 50%.

(b) How is the Pooled Standard Deviation is computed

The pooled standard deviation was calculated for all <u>time 1 - time 2</u> scores in the database. This means that the change scores are standardized across all time periods and all types of children.

(c) Thematic Effect Sizes

The thematic effect sizes were created by averaging component effect sizes.

1. **Positive effect** = attachment + positive peer relations + prosocial behaviour + self efficacy + social support + quality of life
   - Tests combined for the mean positive effect are:
     - Kern's Security Scale, total score
     - Parental Bonding Instrument, caring and over-control scales
     - IPPA total score
     - Child BASC parent-child relationship scale
     - Piers Harris Happiness scale
     - SDQ – prosocial scale
     - Self efficacy scale

---

[9] Jacob Cohen (1988). *Statistical Power Analysis for the Behavioral Sciences* (second ed.). Lawrence Erlbaum Associates

[10] Cohen gives the following guidelines for the social sciences: small effect size, r = 0.1 − 0.23; medium, r = 0.24 − 0.36; large, r = 0.37 or larger

- o Social support scale
- o Overall rating of facility and composite score of client satisfaction survey

2. **Quality of Life for the Medically Fragile** = Health of Skin and Feet + Cognition and Perception + Use of Hospitalization (tests have the same name)
3. **Functioning** = academic performance + acceptance and success fitting into the social environment + independence + adaptive functioning
   - Tests combined for the mean functioning effect are:
     - o Report cards and/or the WIAT for math, reading and oral learning
     - o Children's Global Assessment Scale
     - o Level of Care – total score
     - o ABS, adaptive functioning scale
     - o SDQ – peer problems scale
     - o Piers-Harris total score
     - o Child BASC: self-esteem, self-reliance, locus of control and adjustment
     - o Parent BASC: adaptive, social, daily living and communication skills
4. **Risk** = mental illness + substance abuse and suicide + physical health + ADHD + affect regulation + psychiatric symptoms + developmental anomalies + risk of skin breakdown
   - Tests combined for the mean risk effect are:
     - o Strengths and Difficulties Questionnaire (total problem scale)
     - o GAIN-SS
     - o Behavioural Determinants of Health
     - o Conners' Global Index
     - o Feelings, Attitudes and Behaviour (FAB-C) – total problem Index
     - o SA-45 (Global Symptom index)
     - o Nippissing District Developmental Screen (NDDS)
     - o Braden Scale
5. **Impairment** = serious behaviour problems + dangerous or self harming behaviour + alcohol and drug abuse + neurological symptoms + autism symptoms
   - Tests combined for the man impairment effect are:
     - o Serious Behaviour Problems Checklist
     - o Danger to others Checklist
     - o Harm to Self Checklist
     - o Alcoholism Screen
     - o Drug Use Scale
     - o Neuro-Behavioural Problems
     - o Childhood Autism Rating Scale
6. **Family Life** = family functioning + coping ability when under stress + parental distress + ability to forgive self and others + parental self-efficacy + parental social support
   - Tests combined for the mean family effect are:
     - o Family Assessment Device (General Functioning Scale)
     - o F-COPES
     - o QRS-F
     - o Heartland Forgiveness Scale

- o Self Efficacy checklist
- o Social Support Checklist

Robert Fulton

March 17, 2015