# **Introduction**

…

## **The present study**

Available anonymous Routine Outcome Monitoring (ROM) data, collected as part of MSTs quality assurance and improvement system, was used to investigate the effectiveness of Multisystemic Therapy on treatment outcome during the COVID-19 restriction measures in Norway compared to before and after. The system has been designed to increase the likelihood that problems with treatment fidelity and treatment outcome are identified and addressed on an ongoing basis at family, therapist, supervisor, expert consultant, and organization operating the treatment level (Henggeler et al., 2009, p. 284). ROM data were reviewed for families starting treatment two years before the national lockdown until about 17 months after. I was thus able to compare the effectiveness of MST for families who received treatment when the lockdown occurred and those who started treatment after the lockdown date (some of these probably experienced later periods of restrictions) with the families who received the treatment before the pandemic onset. This is an important extension of previous research which has primarily studied MST under normal circumstances where physical meetings are feasible. To fill this gap in the literature, I examine three key questions as part om MSTs continuous quality improvement system: 1) Do treatment outcomes of MST change for those adolescents who receive treatment during or after the COVID-19 lockdown date; 2) Is there a change in which covariates influences the treatment outcome; 3) Is there any differences in the growth patterns for those receiving treatment before, during and after the COVID-19 lockdown date?

To answer these questions, a longitudinal natural-group comparison was used. Given the limited evidence on the effectiveness of interventions for adolescents with serious conduct problems during the COVID-19 pandemic, and on how the situation affected this group of young people, it is difficult to establish testable hypothesis related to these research questions. (også nevne blandet evidens fra ikke-kliniske utvalg?) The present study thus represents an exploratory multi-group comparison of how MST worked during the COVID-19 restriction measures.

# **Methods**

## **Participants**

The participants are MST cases registered in the Norwegian Center for Child Behavioral Development (NUBU) database system, which started treatment between the 13th of March 2018 and the 31st of August 2021. A small portion of this sample have been included in other studies (e.g., Hukkelberg et al., 2022; Keles et al., 2021). A total of 2,067 young people and their families, referred to MST by the municipal Child Welfare Services for serious and persistent conduct problems, received treatment during this period. Due to this studies exclusion criteria (cases with several missing values on YLS/CMI and the national outcome goals at admission to treatment (T0)), 46 cases (2.23%) were excluded from the sample. Five of these cases had missing values on both YLS/CMI and the national outcome goals, and 11 had missing on the latter. While 30 had missing on all variables except from treatment duration and region, due to the family`s lack of consent to register case data. This resulted in a final sample of 2021 young people (63.4% boys) between 7 and 18 years, with a mean age of 14.33 years (add .05 years) (*SD* = 1.60). 80.2% of the adolescents were non-immigrants, 10.8% were immigrants, and 9.0% had immigrant parents. Of those with an immigrant background, the majority were of Asian (40.3%), European (28.5%) or African (21.2%) origin. Mean treatment duration for the total sample was … months (SD = …).

In order to receive MST treatment, the youth must meet the following inclusion criteria, evaluated by the team supervisor based on the referral information and the teams initial assessment: 1) the adolescents were between 12 and 18 years (NUBU, 2022b), 2) the adolescent displayed serious rule or norm breaking behavior, 3) the adolescent is at immediate risk of out-of-home placement (NUBU, 2021). While youths with the following characteristics are to be excluded from the intervention: 1) adolescents living by themselves and do not have any adults who can act as parents or primary caregivers, 2) adolescents with serious mental health problems (e.g., is actively suicidal, is psychotic, or is a danger to themselves or others), 3) sexual offending youths (without other criminal or antisocial behavior), 4) the adolescent struggles with social communication and interaction, and repetitive behavior, which may be caused by an autism spectrum disorder, 5) the adolescents level of intellectual ability is the most direct contributor to the referral behavior. Despite these inclusion criteria, 117 young people under the age of twelve is admitted to treatment. This only happen exceptionally, but is usually due to an older sibling in the family already receiving MST treatment.

## **Data Collection Procedures**

Under the MST protocol, data on [behavioural 5] and YLS shall be collected at admission (*T0*) and discharge (*T1*). The MST team then continues with data collection on [behavioural 5] at 6-, 12-, and 18-month post discharges (*T2* to *T4*) by phone interviews with youth’s parents. To enhance measurement quality, a client’s condition at admission is co-assessed by the referral agencies’ documentation and by the MST team’s initial clinical interview. During treatment, a quality assurance employee maintains monthly fidelity assessment with youth’s primary caregivers by phone. All MST therapists and agents have undergone rigorous training and must follow detailed guidelines throughout the engagement period.

## **Measures**

All measures except *the Youth Level of Service/Case Management Inventory* (Hoge & Andrews, 2011) have been developed as part of the quality assurance system of MST. Communication channels were designed to be responsive and efficient to ensure immediate feedback between therapists and MST administrators, with minimal written text demanded from frontline agents on top of legally required paperwork (Henggeler et al., 2009, p. 18 – 19, 42). Reliability coefficients of these measures were not made available to the author.

This study partitions Norway into southern, middle, and northern regions to account for geographical variations. All dichotomous variables are scored with (1) *yes* and (0) *no*, such as to questions a) whether the young person lives in foster care, b) has received or is receiving mental health services, and c) has been returned from institutional treatment to receive MST. Demographic information such as sex and age was recorded at admission. Clients born to at least one Norway-born parent were categorized as non-migrant, those born in Norway to both foreign-born parents were categorized as second-generation migrants, and those born overseas as first-generation migrants (NUBU, 2021). Prior to MST admission, clients may have received one or more therapies such as Functional Family Therapy, Treatment Foster Care Oregon, Parent Management Training Oregon, Multifunctional Treatment in Residential and Community Settings, and MST itself. I re-coded “having received prior therapies” into a dichotomous response. Lastly, treatment duration was calculated by subtracting the date of discharge (*T1*) from the date of admission (*T0*), then converted to months. *T2* to *T4* was constructed by adding 6 months to each previous time point.

### *National outcome goals*

Five overarching behavioral goals for MST treatment were assessed at admission and discharge for youth and their families who not completed treatment, and at all time points (i.e., admission, discharge, and 6-, 12- and 18-month follow-ups) for those who completed treatment (NUBU, 2021). These ultimate outcomes are: 1) lives at home (i.e., the youth was not in jail or placed out of home by the Child Welfare Services), 2) attends school or work (at least 50% workload), 3) law-abiding (i.e., the youth has not received any formalized, societal consequences as a result of offenses or misdemeanors), 4) drug-free (i.e., the youth did not use drugs in a way that affects daily functioning or led to other serious consequences), and 5) abstaining from violence/ threats of violence. Non-risk outcomes were rated as (1) *yes*, and risk outcomes as (0) *no*.

### *Risk level*

The level of youth risk for future conduct problems was measured by the Youth Level of Service/Case Management Inventory – Part I (YLS/CMI; Hoge & Andrews, 2011), both at intake to treatment and at discharge. YLS/CMI consists of 42 items summarized across eight subscales, rated on a dichotomous scale (1 = present, 0 = not present). The first subscale Prior and Current offenses/ Dispositions (5 items), measures static risk (i.e., unable to be changed). While dynamic risk (i.e., able to be changed) is measured by the remaining subscales; Family Circumstances/ Parenting (6 items), Education/ Employment (7 items), Peer Relations (4 items), Substance Abuse (5 items), Leisure/ Recreation (3 items), Personality/ Behavior (7 items), and Attitudes/ Orientation (5 items). Both a total scale score and subscale scores are calculated, where a higher score indicates greater risk. The Norwegian version of the YLS/CMI was translated by Tore Andreassen and Terje Ogden, and back-translated to gain approval by the Multi Health Systems Inc. in the US (Jakobsen & Kornør, 2017). In the present study, I had access to subscale data only, not item data. Consequently, I was not able to calculate the internal consistency. However, in their study on parts of the same sample, Hukkelberg et al (2022) report an internal consistency of α = .55 at intake and α = .71 at discharge.

### *Treatment fidelity*

The therapist adherence to the MST treatment model and the nine treatment principles was measured using *the Therapist Adherence Measure - Revised* (TAM-R; Henggeler et al., 2006), a questionnaire of 28 items scored in a 5-point Likert scale ranging from (1) *not at all* to (5) *very much*. A total TAM-R score is calculated by dividing the number of items answered with *very much* by the total number of items answered, resulting in a score ranging from 0 to 1. Cut-off for treatment fidelity has been set to 0.61 (NUBU, 2021).

## **Statistical analysis**

Analyses were performed using Mplus 8.9 (Muthén & Muthén, 2017). To compare the effectiveness of MST treatment before, during and after the lockdown date, I divided the sample into three groups: 1) before (all cases finishing treatment before the 12th of March 2020), 2) during (all cases that receive treatment on 12th of March), and 3) after lockdown (all cases that received the full treatment phase after the 12th of March). This grouping variable serves as the independent variable in the analysis. Throughout the rest of this article, I will talk about these groups and not the overall sample.

In answer to the first and second question I am evaluating the effectiveness of the lockdown on the outcome of MST treatment using two linear models. The number of achieved ultimate outcomes at discharge (generated by collapsing the separate national outcome goals into one continuous variable, ranging from 0 - 5) is the dependent variable in one of these models, while the other model uses the YLS/CMI total at discharge. Due to this study’s exploratory design, several variables were included as covariates in the linear models. Including TAM-R and demographic information such as gender (a dummy variable, with male as reference level), age, lives in foster care, received health care, returned from institutional treatment, a dummy coded earlier treatment variable (the different previous treatments collapsed into one prior treatment category (1) vs. no prior treatment (0)), two immigrant status dummy variables (immigrants (named “immi1”) vs. nonimmigrants and immigrant parents (0); immigrant parents (named “immi2”) vs. nonimmigrants and immigrants (0)), and two dummy coded treatment region variables (middle (1) vs. the 4 other regions (0); north (1) vs. the 4 other regions (0)). In addition, baseline scores on the ultimate outcomes were added to the ultimate outcome model as a covariate, and the baseline YLS/CMI total to the YLS/CMI model. In this way, I control for potential differences at intake between the groups. By doing so, three results tables with main effects for each of the two linear models are generated, one for each group. Specifically, to answer the first research question about whether there are any changes in treatment outcomes I investigate the intercept of the linear models. In addition to … (). The standardized slope coefficients in the two models were compared separately across groups to answer the second question about changes in covariates.

In answer to the third research question about differences in growth patterns for the tree groups responses to MST, I apply a one faced latent growth curve model (LGCM). This kind of statistical approach is widely used for analyzing change over time (Geiser et al., 2013). The LGCM constitute a multi-group analysis, due to the inclusion of the grouping variable, with five time-points (T0-T5). Data collected at admission (T0), discharge (T1), and 6- (T2), 12- (T3), and 18- (T4) month follow-ups after discharge on the number of achieved ultimate outcomes were used as dependent variables in the analysis. The discharge time (T1) is set to vary by number of months of treatment the youths received (the interval between T0 and T1 thus vary by youths), while T0 and T2-T4 has a fixed time interval (T0 = 0 months, T2 = 6 months after discharge, T3 = 6 months after T2, and T4 = 6 months after T3). By doing so, the LGCM obtain an exact growth trajectory of each adolescent. (Tony – do I have to report the factor loadings?). The covariates included in the growth curve model are the same as in the linear models, except from the baseline scores on the outcome variables which is excluded. Instead, a change score on YLS/CMI total is included, calculated by subtracting the baseline score by the discharge score. The model included an intercept and a linear slope, instead of a quadratic slope. This decision was made due to the testing of nonlinearity in individual lines of trajectory, where every indicator was non-significant (Tony – Do we have this documented?). This indicates that the data fit the linear model better than the nonlinear model. In this unstandardized model I only interpret the direction of the coefficient, not the quantity, due to the covariates different scales, which makes it difficult to compare the coeffisients of the different covariates (Navarro & Foxcroft, 2022, p. 307 - 308).

For all three models I employed Mplus`s restricted maximum likelihood estimator (MLR; Muthén & Muthén, 2017) thanks to MLR`s robustness against potential violation of Gauss-Markov assumptions (Wooldridge, 2020).

### *Missing data and attrition*

The percentage of missing values across the 19 variables varied between 0 and 50.22%. In total … out of … records (…%) were incomplete. Many families had no TAM-R score due to the European Union`s legislation (Schrems II; Datilsynet, 2021) which came in force the summer 2020, regulating the transfer of personal data to countries outside the European Economic Area (EEA). Due to ethical considerations, no follow-up information has been collected for non-completers, which may partly explain the amount of missing data on these variables. I used multiple imputation to create and analyze 10 multiply imputed datasets. Multiple imputation is regarded as the best practice for handling missing data because it improves statistical power and accuracy relative to other techniques (van Buuren, 2018). Incomplete variables were imputed using Mplus`s internal imputation capacity (Muthén & Muthén, 2017). The parameters were estimated in each imputed dataset separately, and combined using Rubin`s rules (**this is copy-paste**). For comparison, I also performed the analysis on the subset of cases without missing data at the independent variables, which hardly affect the parameters (Tony - is this the correct formulation?)

# **Results**

## **Descriptive statistics**

Table 1 presents the descriptive statistics for both the overall sample, and the before, during and after lockdown groups. This table shows, for example, that 40.8% of those who received treatment when Norway introduced the national lockdown were girls. (**Place Table 1 here**).