Data Supplement for McNiel and Binder, Effectiveness of a Mental Health Court in Reducing Criminal Recidivism and Violence, American Journal of Psychiatry, September 2007

Supplemental Material on the Methodology and Results of the Study

Here we discuss subsidiary analyses with the subset of the sample with schizophrenia, and details of our method of calculation of the average treatment effect.

Part A. Subsidiary Analyses of Persons with Schizophrenia

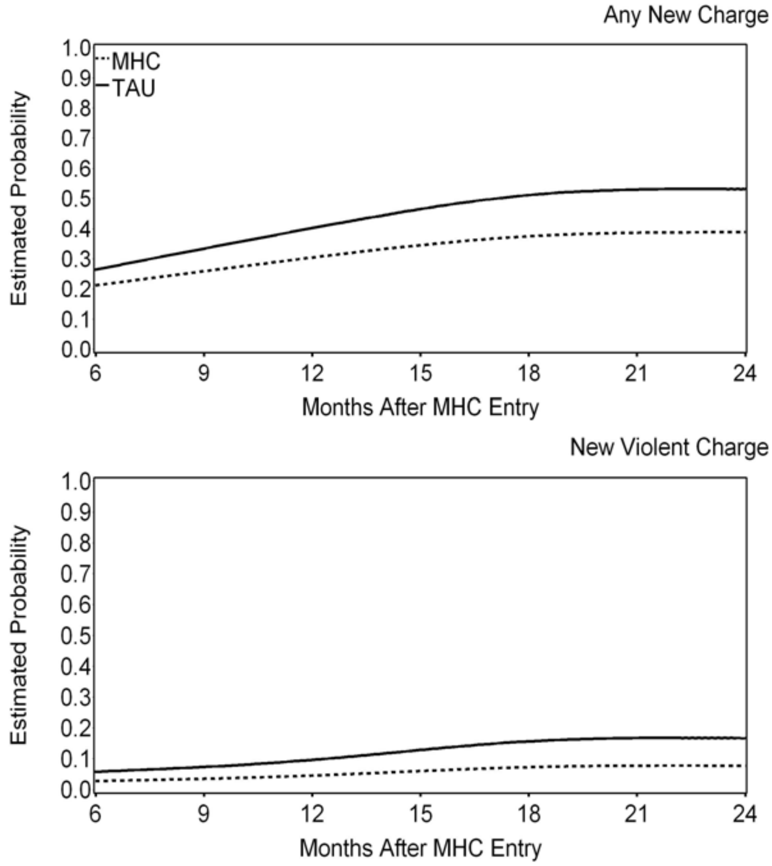
As an additional check on our methods, we performed subsidiary analyses that were restricted to persons with schizophrenia (66% [112/170] of the mental health court participants and 7% [549/8,067] of the treatment as usual group]. We performed intent-to-treat analyses that compared the mental health court participants and the treatment as usual group with schizophrenia from the time of enrollment to the end of the follow-up period, without distinguishing whether individuals successfully completed the mental health court program. After controlling for propensity scores, demographic characteristics (e.g., sex and age), comorbid diagnoses, and number and type of charges during a 12-month baseline period, Cox proportional hazard models showed that mental health court participation predicted a longer time to any new charge ($B = –0.38$, $p<0.01$) and longer time to a new violent charge ($B = –0.058$, $p<0.05$). Supplemental Figure 1 plots the average probability of a new charge as a function of months after entry to mental health court or treatment as usual based on the Cox model. (Supplemental Table 1 summarizes the plots at 6-month intervals and reports the 95% confidence intervals for the effect of mental health court participation on reducing the probability of a new charge.) Supplemental Figure 1 shows that the effect of mental health court participation becomes more evident over time. By 18 months after entry, the cumulative probability of any new charge for the mental health court participants and the treatment as usual group, respectively, was 0.38 and 0.51; at 18 months, the average reduction in the probability of a new charge due to mental health court participation was –0.14 (95% CI = –0.16 to –0.11), which represents a 27% (0.14/0.51) reduction. Similarly, Supplemental Figure 1 shows that the cumulative probability of a new violent charge, for mental health court participants and the treatment as usual group, respectively, was 0.07 and 0.16, at 18 months; the average effect of mental health court participation on reducing the probability of a new violent charge was –0.08 (95% CI = –0.09 to –0.08), which represents a 50% (0.08/0.16) reduction.

Supplemental TABLE 1. Average Effect of Mental Health Court Participation on the Probability of a New Charge for Persons With Schizophrenia

Months of Follow-Up	Mean (SD) Probability of New Charge		Average MHC Effect (SD): MHC – TAU	Standard Error of the Average MHC Effect	t	95% CI for the Average MHC Effect
Months After Entry	MHC[a]	TAU[a]				
Any New Charge						
6 Months	0.21 (0.20)	0.27 (0.07)	–0.05 (0.18)	0.023	–2.28	–0.10 , –0.01
12 Months	0.30 (0.25)	0.40 (0.09)	–0.10 (0.22)	0.011	–9.13	–0.12, –0.08
18 Months	0.38 (0.28)	0.51 (0.09)	–0.14 (0.25)	0.010	–13.88	–0.16 , –0.12
24 Months	0.39 (0.28)	0.53 (0.09)	–0.14 (0.25)	0.010	–14.69	–0.16 , –0.13
New Violent Charge						
6 Months	0.03 (0.06)	0.06 (0.06)	–0.03 (0.05)	0.002	–16.01	–0.03 , –0.03
12 Months	0.05 (0.08)	0.10 (0.08)	–0.05 (0.07)	0.003	–18.35	–0.06 , –0.05
18 Months	0.07 (0.12)	0.16 (0.11)	–0.08 (0.10)	0.004	–21.58	–0.09 , –0.08
24 Months	0.08 (0.12)	0.17 (0.11)	–0.09 (0.11)	0.004	–21.98	–0.10 , –0.08

[a] MHC = mental health court; TAU = treatment as usual.

Supplemental FIGURE 1. Estimated Cumulative Probability of a New Charge for Persons With Schizophrenia by Mental Health Court Status and Months After Entry[a]



[a] MHC = mental health court; TAU = treatment as usual. The estimates are calculated from the propensity-weighted Cox proportional hazard model results.

In the aggregate, these results show similar average effects of mental health court participation on reduction in recidivism when analyses were restricted to persons with schizophrenia, as were evident in the more comprehensive analyses with the full study group that are reported in the print version of this manuscript.

Part B: Method of Calculating the Average Treatment Effect

B.1. Average Treatment Effects for the Cox proportional Hazards Model

Let us define *MHC* as an indicator variable for mental health court participation, where

$$MHC = \begin{cases} 0, & \text{if a TAU participant} \\ 1, & \text{if a MHC participant} \end{cases},$$

and let *X* be a *k*-dimensional vector of pretreatment variables. Assuming proportional hazards and a linear regression function for the pretreatment variables *X*, the logarithm of the hazard function for the Cox regression model is

$$\log\left(h(t, MHC, X)\right) = \alpha MHC + \sum_{j=1}^{k} \beta_j X_j + \sum_{j=1}^{k} \Delta_j X_j MHC$$

where the indicator variable *MHC*, each pretreatment variable X_j, and the product of *MHC* and each X_j are included as covariates.

For the subsample with the pretreatment variables $X = x$, with *x* being a particular set of values for the elements of *X*, the so-called *survival function* associated with the Cox model is defined by

$$S\left(t, x, MHC\right) = S_0\left(t\right)^{\exp\left(\alpha MHC + \sum_{j=1}^{k} \beta_j x_j + \sum_{j=1}^{k} \Delta_j x_j MHC\right)},$$

and the failure function by:

$$F\left(t, x, MHC\right) = 1 - S\left(t, x, MHC\right) = 1 - S_0\left(t\right)^{\exp\left(\alpha MHC + \sum_{j=1}^{k} \beta_j x_j + \sum_{j=1}^{k} \Delta_j x_j MHC\right)},$$

where $S_0\left(t\right)$, the baseline survival function, is given by $-\log\left(S_0\left(t\right)\right) = H_0\left(t\right)$, the baseline cumulative hazard function (1).

Part B.2. Average Treatment Effect on the Failure Function

We are interested in the *average treatment effect* on the *failure function*, where the *average treatment effect* is taken to mean the amount by which the probability of reoffending as a function of time since baseline differs *on average* between two counterfactual scenarios—one where all individuals would be assigned to mental health court and the other where all individuals would be assigned to treatment as usual, conditional on all the above-mentioned confounding variables included in the Cox regression model. What would the expected treatment effect for an individual be if we were to manipulate the treatment condition for that individual from treatment as usual to mental health court, holding all x_j's fixed? We describe these scenarios as counterfactual, since for those individuals in the sample who were actually in the mental health court group we do not have observations under the treatment as usual condition and, similarly, for those individuals who were actually in the treatment as usual group we do not have observations under the mental health court condition. This counterfactual construct is central to the treatment effect literature (2–6). Hence, the treatment effect for an individual can be formally expressed as

$$\begin{aligned}\tau_F(t, x_i, MHC) &= E\left\{\left[F(t, x_i, MHC_i, \alpha, \beta, \Delta) \middle| MHC_i = 1\right] - \left[F(t, x_i, MHC_i, \alpha, \beta, \Delta) \middle| MHC_i = 0\right]\right\} \\ &= E\left[F(t, x_i, MHC_i, \alpha, \beta, \Delta) \middle| MHC_i = 1\right] - E\left[F(t, x_i, MHC_i, \alpha, \beta, \Delta) \middle| MHC_i = 0\right] \\ &= F_i^{MHC}(t, x, MHC, \alpha, \beta, \Delta) - F_i^{TAU}(t, x, MHC, \alpha, \beta, \Delta).\end{aligned}$$

Clearly, $\tau_F(t, x_i, MHC)$ involves counterfactuals. However, it is also clear that $\tau_F(t, x_i, MHC)$ represents the expected failure for an individual under the two different treatment conditions. Therefore, the *average treatment effect for individuals* is

$$\tau_F(t, x, MHC) = E\left[F^{MHC}(t, x, MHC, \alpha, \beta, \Delta) - F^{TAU}(t, x, MHC, \alpha, \beta, \Delta)\right].$$

Following Terza (2) and Basu and Rathouz (3), once consistent estimates are obtained for the parameters of the Cox model $\left(\hat{\alpha},\, \hat{\beta}_j \text{ and } \hat{\Delta}_j\right)$, $\tau_F\left(t, x, MHC\right)$ can be consistently estimated by the sample average of the individual differences:

$$\hat{\tau}_F\left(t,x,MHC\right)=\frac{1}{n}\sum_{i=1}^{n}\left[\hat{F}_i^{MHC}\left(t,x,MHC,\hat{\alpha},\hat{\beta},\hat{\Delta}\right)-\hat{F}_i^{TAU}\left(t,x,MHC,\hat{\alpha},\hat{\beta},\hat{\Delta}\right)\right],$$

which is the value of $\tau_F\left(t,x,MHC\right)$ that minimizes the objective function

$$\sum_{i=1}^{n}\left\{\left[\hat{F}_i^{MHC}\left(t,x,MHC,\hat{\alpha},\hat{\beta},\hat{\Delta}\right)-\hat{F}_i^{TAU}\left(t,x,MHC,\hat{\alpha},\hat{\beta},\hat{\Delta}\right)\right]-\tau_F\left(t,x,MHC\right)\right\}^2 .$$

The estimate $\hat{\tau}_F\left(t,x,MHC\right)$ depends on the Cox model estimates $\hat{\alpha},\, \hat{\beta}_j$ and $\hat{\Delta}_j$, and thus the usual calculated regression standard error for $\hat{\tau}_F\left(t,x,MHC\right)$ is inconsistent and cannot be used. A consistent estimate of its standard error can be obtained by applying the method of Murphy and Topel (7) for two-step estimation to the objective function (see references 8 and 9 for further description of this method).

Since $\tau_F\left(t,x,MHC\right)$ is the expected value of the failure function due to an experimental manipulation of treatment conditions for each individual, it can be characterized as representing the expected "causal" effect of treatment on the failure function. For complete details on Terza's method, including proofs, see Terza (2). All results for Terza's method hold when a propensity-weighted likelihood for the Cox regression model is used to obtain the estimates $\hat{\alpha},\, \hat{\beta}_j$ and $\hat{\Delta}_j$.

References

1. Kalbfleisch JD, Prentice RL: The Statistical Analysis of Failure Time Data. New York, John Wiley & Sons, 1980
2. Terza JV: Parametric regression and policy analysis: a framework for estimation and inference. Working Paper, Department of Epidemiology and Health Policy Research, University of Florida, 2006
3. Basu A, Rathouz PJ: Estimating marginal and incremental effects on health outcomes using flexible link and variance function models. Biostatistics 2005; 6:93–109
4. Imbens GW: Nonparametric estimation of average treatment effects under exogenicity: a review. Review of Economics and Statistics 2004; 86:4–29
5. Hirano K, Imbens GW: Estimation of causal effects using propensity score weighting: an application to data on right heart catheterization. Health Serv Outcomes Res Methodol 2001; 2:259–278
6. Rosenbaum PR, Rubin DB: The central role of the propensity score in observational studies for causal effects. Biometrika 1983; 70:41–55
7. Murphy KM, Topel RH: Estimation and inference in two-step econometrics models. Journal of Business and Economic Statistics 1985; 3:370–379
8. Greene WH: Econometric Analysis, 3rd ed. New York, Prentice Hall, 1997
9. White H: Estimation, Inference, and Specification Analysis. New York, Cambridge University Press, 1994