Comment: OCTET does not prove community treatment orders are ineffective

The original published article is available from The Lancet Psychiatry

The OCTET randomised controlled trial (RCT) of supervised community treatment orders (CTOs) for patients with psychosis found no difference in the primary outcome of readmission to hospital in those with CTOs compared with those receiving section 17 leave under the England and Wales Mental Health Act, which provides similar levels of compulsory supervision and treatment before final discharge from hospital. Despite the authors’ conclusions that CTOs do not benefit patients, they have continued to be widely used.2

As a result of legal necessity, OCTET randomised two different types of compulsory treatment, with the expectation that those assigned to section 17 would very quickly be discharged to a voluntary status. However, in practice, patients were often subject to community compulsion and treatment for many weeks with a mean of 45 days (median 8 days). Evidence shows that CTOs might be effective but only over durations longer than 6 months.3,4 Since an average admission rate of once every 2 years occurred in OCTET recruits, patients might not normally be expected to be readmitted to hospital within the mean CTO duration time of 170 days (median 183 days). As a result, the question is raised as to whether enough time was available to enable any benefits to become apparent.

The length of time under compulsion might represent a normal outcome of CTOs, rather than a limitation of the trial itself.5 However, since 2011, about 4000 new CTOs have been given per year in England and Wales,6 in which the prevalence of CTOs has been roughly 5000,2 showing an average duration of 14 months (ie, the point prevalence divided by annual incidence). These data suggest that the short duration of CTOs in OCTET might have been related to the mechanics of the trial, rather than an outcome which would be expected in relation to their routine use.

Questions have also been raised as to whether the participants in OCTET were those most likely to benefit, partly based on the 20% refusal rate, and partly as a result of the possibility that the recruiting psychiatrists only entered people who they were less sure would need a CTO compared with others who they might have chosen not to put forward for the trial.7 These issues, although subjective and unprovable, might undermine OCTET’s conclusions.

To conclude that a trial has proved an intervention does not work, the trial must be adequately powered until completion. OCTET assumed an effect size of 16% and was powered at 84.7% at the start. However, with 24% of people not receiving their assigned intervention or the opposite intervention, we show with a deterministic method, and a Monte Carlo simulation that power fell to 35.2%, which is well below that required to prove no type 2 error is present (appendix). Hence, OCTET has only a 35.2% chance of proving the effectiveness of CTOs, and seven negative OCTET trials would be needed to be 95% certain of a true negative result.

So where does this leave the clinician wanting to take an evidence-based approach to deciding whether to use a CTO? The New York RCT4 was also likely to have been underpowered, as a result of a very high attrition rate. Additionally, the conclusions of the North Carolina study,3 which was the only other RCT that showed compulsory supervision and treatment-conferred benefits when it was provided for more than 6 months, have been questioned due to methodological concerns. Although evidence for the benefits of CTOs exists from observational studies, results are
inconsistent.\textsuperscript{3} As a result, CTOs are in the category of many legal interventions, such as compulsory admission to hospital, that are not disproved but lack robust evidence to prove effectiveness.

New or improved evidence from RCTs is unlikely to become available in the foreseeable future and thus, use of available evidence is essential. A more detailed investigation into the per-protocol analysis of days spent in hospital might therefore be justifiable. Although this investigation introduces potential bias, the tendency for intention-to-treat trials to increase the risk of type 2 error has to be balanced. Without adjustment for crossovers, OCTET showed a non-significant 13\% reduction in nights spent in hospital for people on a CTO compared with those in the section 17 group, when patients who were never discharged were excluded. A per protocol analysis\textsuperscript{8} was significant (p=0.035) but the means and standard deviations have never been published. With the assumption of similar standard deviations as the original analysis, we estimated the effect size to be around a 16–25\% reduction in admission length. Hence, the per-protocol analysis lends support to the original but underpowered analysis, showing similar effects which are significant and clinically important in size. This analysis does not prove CTOs are effective but provides a useful addition to the existing evidence.

We suggest that OCTET cannot claim to have disproved the effectiveness of CTOs. The trial might suggest hospital time for CTOs was reduced but we question the original conclusions and associated calls for a moratorium on their use.\textsuperscript{9}

\textsuperscript{*}Andy Owen, Loopinder Sood

Coventry and Warwickshire NHS Partnership Trust, Stratford Healthcare Building 2, Arden Street, Stratford upon Avon, CV37 6NQ, UK (AO); and Coventry and Warwickshire NHS Partnership Trust, St Michaels Hospital, St Michaels Rd, Warwick, CV34 5QW, UK (LS)

andy.owen@covwarkpt.nhs.uk

We declare no competing interests.


Appendix: an assessment of power in the OCTET trial by Monte Carlo simulation and by deterministic methods

In OCTET 24% of the section 17 group were placed on a CTO, and 21% of the CTO group were discharged without a CTO in place. Thus, subjects crossed over to the opposite treatment limb which we call “crossovers”. Additionally almost 4% were randomised but never left hospital. These protocol violations greatly affect power. It is possible to calculate this effect on power using the deterministic method outlined by Wittes. Crossovers reduced the effect size from 16-0% (48-0%-32-0%) to 8-4% (42-4%-34-0%) (box 1).

\[
\begin{align*}
\pi_r &= \pi_{ct} + (1 - \pi_{ct} \cdot \pi_{ct}^*) \pi_t \\
\pi_r^* &= (1 - \pi_{ct} \cdot \pi_{ct}^*) \pi_{ct} + \pi_{ct} \pi_t
\end{align*}
\]

where,

\(\pi_t\) and \(\pi_{ct}\) = readmission rates treatment and control groups

\(\pi_{t}^*\) and \(\pi_{ct}^*\) altered readmission rates in = treatment and control groups after crossovers and dropouts

\(\pi_t\) = proportion dropped out of treatment and crossover to control group

\(\pi_{ct}\) = proportion never discharged in treatment group

\(\pi_{ct}^*\) = proportion who dropped into treatment from control group

\(\pi_t^*\) = proportion never discharged in control group

This reduced effect size was used in the standard formula for obtaining power when testing the difference between two proportions, which showed reduction in power from 84-7% to 35-2%. It was found that to maintain power at 80%, with this reduced effect size, the total sample size would need to be 1044, over 3 times larger than in the OCTET trial.

It is possible to determine whether the lack of observed difference found in the trial was significant, by Monte Carlo simulation.

The probability of readmission on which OCTET’s power calculation was based was 0.32 in the treatment group and 0.48 in the control group. With no crossovers the number of admissions in the control group was obtained by generating 167 uniformly distributed random numbers in the range zero to one. Numbers thus generated which were less than 0.48 were counted as an admission in the control group. Similarly, in the treatment group, any of another set of 166 random numbers less than 0.32 were counted as an admission. To simulate the effect of a certain percent of crossovers, that percent of generated random numbers were counted as an admission if less than the admission probability of the opposite treatment limb. To simulate the effect of subjects who were never discharged the same amount of random numbers were counted as a zero admission whatever their magnitude.

The OCTET trial was simulated 400,000 times by repeating the above procedure by computer using our Monte Carlo Simulator program.

The frequency of a significant difference in the number of admissions between the two groups, with crossovers and dropouts or without, gave the power. With no crossovers or dropouts, power was 84-5%, consistent with the intended power of the study. With crossovers and dropouts as in OCTET, power fell to 34-7%, consistent with the deterministic method outlined earlier.

The Monte Carlo simulation showed the difference in the number of admissions (control group minus CTO group) was between...
-3 to 31, with 95% frequency (therefore giving the 95% confidence interval). Thus the observed difference of one in OCTET which falls inside this interval does not mean that the study showed no difference when there are crossovers and dropouts, as such a small difference can still arise even when there is an actual difference of 16%. The simulation also showed that with no crossovers and dropouts the 95% confidence interval for this difference was 10 to 44 and if the observed difference had still been one then it would have been well outside this interval and only then it would have meant that the trial showed no difference. Similar results were obtained deterministically by first calculating the expected 95% confidence interval for the difference of proportions and then multiplying by 166 to give the 95% confidence interval for the difference of readmissions of between -4 to 30 with crossovers and dropouts and 10 to 44 without them.

Hence two different types of method lead to the same conclusion that OCTET was underpowered and the observed difference of just one in readmissions does not mean that the result has been shown to be negative with any degree of confidence.

If CTOs were effective, with an effect size of 16%, and having proved that OCTET was powered at 35%, we calculated how many negative OCTET trials would be needed in order to be 95% certain that the result was truly negative using the cumulative function of the geometric distribution (table 1).

**Table 1: probability of a positive result after n studies powered at 35% with 16% effect size**

<table>
<thead>
<tr>
<th>n</th>
<th>Probability of positive result</th>
<th>Cumulative Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.35</td>
<td>0.35</td>
</tr>
<tr>
<td>2</td>
<td>0.2275</td>
<td>0.5775</td>
</tr>
<tr>
<td>3</td>
<td>0.1479</td>
<td>0.7254</td>
</tr>
<tr>
<td>4</td>
<td>0.0916</td>
<td>0.8215</td>
</tr>
<tr>
<td>5</td>
<td>0.0625</td>
<td>0.884</td>
</tr>
<tr>
<td>6</td>
<td>0.0461</td>
<td>0.925</td>
</tr>
<tr>
<td>7</td>
<td>0.0264</td>
<td>0.951*</td>
</tr>
</tbody>
</table>

*reaches 95% significance level

The table shows that with power reduced to 35% it would take 7 negative trials to be 95% sure that the result is truly negative. This is because the cumulative function of the geometric distribution tells us that if the treatment effect was actually there then it would have been picked up by at least one of the seven trials with probability of 95%.

**Reference**
