Duration of therapy – Does it matter?☆
A systematic review and meta-regression of the duration of psychosocial treatments for alcohol use disorder

Lotte Kramer Schmidt a,⁎, Anders Bo Bojesen b, Anette Søgaard Nielsen b,c, Kjeld Andersen b,d

a Unit of Clinical Alcohol Research, University of Southern Denmark, J.B. Winsløwsgade 20, entrance 2208, 5000 Odense C, Denmark
b Unit of Clinical Alcohol Research, University of Southern Denmark, Denmark
c OPEN Odense Patient data Explorative Network, Denmark
d Department of Mental Health, Region of Southern Denmark, Denmark

A R T I C L E   I N F O
Article history:
Received 16 June 2017
Received in revised form 26 October 2017
Accepted 6 November 2017

Keywords:
Alcohol use disorder
Duration
Dose-effect
Psychotherapy
Psychosocial treatment
Research assessments

A B S T R A C T
Background: The recommendations in clinical guidelines for duration of therapy for alcohol use disorder (AUD) are based on consensus decisions. In reality, we do not know the optimal duration of an alcohol treatment course. Methods: A systematic review and meta-regression of randomized controlled trials of psychosocial treatment in alcohol outpatient treatment centers. The population consisted of adults suffering from AUD, treated in an outpatient facility with at least two sessions of therapy. Meta-regression analysis was performed with treatment outcome as a function of duration of therapy across studies. Treatment outcome was defined as long-term alcohol use measured in percentage of days abstinent (PDA), percentage of heavy days drinking (PHD), and/or proportion of participants abstinent (ABS).
Results: 48 studies encompassing 8984 participants. Mean planned duration of therapy: 18 (8–82) weeks and 14 (2–36) sessions. Mean actual attended sessions: 9 (1–26). Mean follow-up time: 43 (8–104) weeks with a mean of 6 (2–18) research assessments. Neither planned weeks, duration of sessions, frequency of sessions per week, nor actual attended sessions were associated with long-term alcohol use outcomes. However, frequency of research assessments was positively associated with PDA and PHD.
Conclusion: No associations between long-term alcohol outcomes and planned or actual attended duration of psychosocial treatment in outpatient care. Research assessments and, accordingly, the research project in itself may influence outcome in studies of psychosocial treatment for alcohol use disorder.
© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Many different psychosocial treatments are offered to patients with alcohol use disorder (AUD). In large part, they have been found equally effective (Martin & Rehm, 2012). While some therapies are conducted over a few weeks, others may last for years. But what is the optimal duration of therapy?

The question about duration of therapy is not new. Dose-response research in psychotherapy in general has sought to answer the question: “How much therapy is enough?” Two major models of how to explain and study the associations between duration and outcome are the “dose-effect” model and the “good-enough level” model. The dose-effect model is based on a medical understanding of dose and assumes a positive association between outcome and dose in the form of sessions demonstrating a negatively accelerating curve: that is, patients improve as the number of sessions increases, but at higher doses the benefit of additional sessions decreases (Kopta, 2003). Based on the dose-effect model, reviews of the duration of psychotherapy estimate that after 13 to 18 sessions, 50% of the patients achieve a good clinical outcome (Hansen, Lambert, & Forman, 2002; Howard, Kopta, Krause, & Orlinsky, 1986). The good-enough level model is based on the belief that patients respond to treatment at different rates and that outcome trajectories are steeper for patients attending fewer sessions (Barkham et al., 2006). This indicates that longer treatment duration might be associated with less rapid rates of change at the individual level (Barkham et al., 2006). Reviews of duration of therapy for substance use disorder have applied the dose-effect model, but have focused on planned duration of continuing care. Continuing care is defined as treatment after intensive...
in- or outpatient care (Lenaerts et al., 2014). Small to no advantages for longer planned durations of continuing care have been found (Blodgett, Maisel, Fuh, Wilbourne, & Finney, 2014; Lenaerts et al., 2014; McKay, 2005, 2009). Moreover, motivational enhancement therapy, which often includes four planned sessions of therapy, has been proven as effective a treatment as cognitive behavioral therapies with longer durations (Martin & Rehm, 2012; Smedslund et al., 2011).

In the treatment of substance use disorder, duration and intensity of treatment have been studied in non-comparable ways and with diverging findings. Some have found inverse associations between longer duration of treatment and drug use outcomes. Magill and Ray (2009) analyzed the effect of cognitive behavioral therapy for substance use disorder in 53 studies and found effect-size diminished after twenty sessions of therapy. Another study analyzed 34 outpatient psychosocial intervention studies for substance use disorders (excluding alcohol dependence), and found that the number of treatment weeks was negatively correlated with substance use outcomes (Dutra et al., 2008). Finally, a study of planned duration of inpatient drug abuse treatment found better effects (in 628 participants) after six months (comprising 42 sessions) of treatment than in either three or twelve months of treatment (McCuoker et al., 1997). In contrast, a recent Cochrane review (Gates, Sabioni, Copeland, Le Foll, & Gowing, 2016) of 23 psychosocial intervention studies for cannabis use disorder found positive associations between more than four sessions or four weeks of treatment and an effect on use of cannabis, and a non-randomized study by Moos and Moos (2003) encompassing 473 first time treatment seekers with AUD indicated not <27 weeks of in-or outpatient treatment to be effective at one year follow-up (Moos & Moos, 2003). Moreover, there are expert recommendations for treating AUD as a chronic disease, with a continuum of care of possibly longer duration (McKay & Hiller-Sturmhofel, 2011; Willenbring, 2013) or a stepped-care model (Haber, Lintzeris, Proude, & Lopatko, 2009). Regarding intensity or frequency of treatment, another Cochrane review (Lancaster & Stead, 2017) of behavioral interventions for smoking cessation indicated higher intensity treatment may be better than lower, but in general, intensity of treatment was not associated with use of drugs in the above-mentioned studies (Dutra et al., 2008; Moos & Moos, 2003), except for one which found lower intensity to be better (McCusker et al., 1997).

Reviews of duration of therapy in general and for substance use disorders in particular have assumed treatment uniformity and pooled different treatments. Since variation in treatment effect sizes is low, and other possible causes of the efficacy of treatment than the specific treatment methods are considered (Imel, Wampold, Miller, & Fleming, 2008; Miller & Moyers, 2015), the assumption of treatment uniformity will also be applied in this study.

Given the gap in the literature concerning the appropriate duration of therapy for AUD in alcohol outpatient care, the recommendations for the planned duration of therapy for AUD are based on consensus decisions (Group, 2008; Haber et al., 2009; Health, 2011; Kleber & Association, 2006). If the effect of duration of therapy is unknown, risks are that patients will receive either too little or too much therapy, with burdensome consequences for both themselves and society (Cuijpers, Huibers, Ebert, Koole, & Andersson, 2013).

Relatedly, how much therapy is actually received? Reviews of therapy for AUD are primarily based on planned durations rather than actually attended weeks or sessions. Moreover, duration of therapy can be interpreted in different ways and in this study we want to include both the planned and actual duration in weeks, duration of sessions, and frequency of sessions per week. Another concern of psychotherapeutic research for AUD is the effect of research assessments as studied by Clifford and Maisto (2000). Since there is a risk of research assessments having a therapeutic effect, the duration and frequency of these will also be taken into account (Clifford, Maisto, & Davis, 2007).

The treatment duration of outpatient care for AUD in general has to our knowledge not been investigated, and a search in the databases for randomized controlled studies of different lengths of the same treatment for AUD threw up only one study (Kamara & Van Der Hyde, 1997).

It is, therefore, an open question whether there are associations between alcohol use outcomes and duration of therapy for AUD as in psychotherapy in general. Knowledge in this area could optimize treatment in alcohol outpatient centers.

To test whether there are positive associations between outcome and duration of treatment of AUD – applying the dose-effect model and assuming treatment uniformity – the research questions of this study are:

- Are there positive associations between duration of treatment and long-term alcohol use outcomes in a population randomized to different kinds of outpatient psychosocial therapy for AUD? Duration of treatment defined as planned and attended weeks, number of sessions, and frequency of sessions per week.

2. Methods

2.1. Information sources

We searched PubMed and Psych info, covering the years from 1966 to 2016, using the search terms: “Alcoholism”[Mesh], therapy*, treatment*, intervention*, train*, counsel*, course*, program*, coach*, session*, consultat*, guid*, mentor*, interview*, period*, month*, week*, year*, length, sequence*, time, duration*, schedule*, short-term, long-term, outpatient*, out-patient*, ambulatory, ambulant. Filters used were randomized controlled trials and only articles in English were reviewed.

A search through references of the background literature and of included articles was also reviewed for potential studies.

The searches were performed in November 2016.

2.2. Eligibility criteria

Using the PRISMA guidelines for systematic reviews (Moher, Liberati, Tetzlaff, & Altman, 2010), the following study characteristics were included (PICO):

Study population: A population randomized at the start of different psychosocial treatments at an outpatient facility for AUD by DSM-III or IV (APA, 1980, 1994). If DSM criteria were not described in the inclusion, the characteristics of the study population at baseline were checked to see if they were likely to fulfill criteria for AUD. If the populations were described as alcoholics, alcohol abusers, addicts, alcohol dependent, or had a MAST (Michigan Alcohol Screening Test) (Selzer, 1971) score above five, they were included. Comorbidity was accepted, but the primary aim of the study was to investigate AUD first and foremost. Use of medication for treating AUD was accepted. Furthermore, the population had to be adults (> 17 years) and to have received no more than one month of inpatient treatment prior to the intervention studied.

Intervention: Any psychosocial intervention performed in outpatient alcohol care services with personal contact (face-to-face, telephone, computer), group or individual treatment, and involving a minimum of two sessions. If the control condition in the study fulfilled these criteria, it was also registered as an intervention.

Comparison: The treatments were not compared relative to one another within the same study. Instead, data on treatment duration from each intervention were used in the meta-analysis and compared across studies and treatment methods.

Outcome was the most frequently used measures of long-term alcohol use: percent days abstinence (PDA), percent heavy days drinking
(PHD), drinks per drinking day (DDD), and proportion of participants abstinent (ABS).

2.3. Study selection

The searches were screened by the first author, using the eligibility criteria mentioned above. If exclusion was not possible by title or abstract, the articles were downloaded as full-texts. These were read by three of the authors (Schmidt LK, Nielsen AS, Andersen K) and inclusion was decided based on consensus.

2.4. Data collection

Because we studied duration of treatments across studies and treatment types, i.e. assuming treatment uniformity, both the control condition and the intervention of the studies were in this case considered to be interventions. The following data were registered for each intervention in the studies: number of participants, therapeutic method, whether there was a pharmacological intervention together with the psychosocial intervention, number of weeks planned, mean number of actually attended weeks, number of planned sessions, mean of actually attended sessions, duration of the research study, number of research assessments, baseline and follow-up data of outcome (PDA, PHD, DDD and/or ABS), and whether outcome was reported as intention to treat or as completers. Recent studies emphasize outcome in psychosocial functioning (Wilson, Bravo, Pearson, & Wittkiewitz, 2016), therefore, sub-analysis of this outcome was considered and data on psychosocial functioning at follow-up was registered. Finally, demographic and descriptive data were registered.

2.5. Data items

Treatment duration: The planned and actually attended weeks and sessions as described for each individual treatment arm in the study, starting with treatment in the outpatient setting. Moreover, the frequency of sessions per week was calculated. In a few cases, the studies reported attrition rate in percentage terms, hence, the actually attended weeks/sessions were calculated. Since half of the studies reported actually attended sessions, we contacted the authors of the studies not reporting this information with the result that nine studies subsequently provided the information. Actual attended weeks were reported even more seldom and we concluded that it would not be feasible to collect enough information on this variable.

Therapeutic method: In order to check the assumption of treatment uniformity, the interventions were classified according to methods used. If cognitive and behavioral strategies were described as a substantial part of the treatment, the intervention was classified as cognitive behavioral therapy (CBT). Treatment was classified according to therapy applied, if possible. If the intervention was described as a standard care intervention without further information regarding content, it was listed as treatment as usual.

Follow-up time: The number of weeks from baseline to the follow-up where the primary outcome was measured. If outcomes from several follow-ups were listed and none were specified as the primary outcome, the follow-up after 26 weeks was chosen. This was in order to optimize homogeneity between studies and, rather than measuring the outcome immediately after treatment, long-term outcomes were chosen whenever possible. Moreover, a study by Magill and Ray (2009) reported that 26 weeks is one of the more stable outcome points in comparable fields.

Research assessments: Interviews and follow-ups of the participants at baseline, throughout, and following treatment. Defined as a meeting with research staff (either face-to-face or by phone) where data of alcohol use was registered, but without engagement in treatment. The number of research assessments extended up to and including the one where the primary outcome was registered.

Alcohol use outcomes were PDA, DDD, PHD and/or ABS. If data were presented in a way that made it possible to obtain the outcome through calculations, this was done. Because outcomes on psychosocial functioning differed widely in the studies, only the source of the data was registered.

Intention to treat: If the statistics described an intention-to-treat analysis or outcome, this was reported as based on the number of participants at baseline. Otherwise the data were registered in the analysis as based on completers.

2.6. Statistics

Meta-regression analysis was used to describe associations between outcome and the following explained variables: planned weeks, actually attended weeks, planned sessions, actually attended sessions, frequency of sessions per week, follow-up time, number of research assessments during the study, and frequency of research assessments per month (28 days) of the follow-up time. Intercorrelations between explanatory variables were computed and we used the metareg package for Stata (Harbord & Higgins, 2008). Standard error estimates were calculated for each group based on number of patients included and using the normal approximation method for proportions.

In order to control for the diversity of treatments and studies, we applied several binary control variables: 1) whether the intervention was classified as CBT, 2) analysis based on intention to treat, 3) pharmacological intervention as part of the psychosocial intervention, 4) abuse of other drugs in the population, and 5) severe psychiatric comorbidity in the population. Finally, adjustments on other demographic data as described in data collection were considered.

2.7. Publication bias

Since we included the control groups of the studies as interventions and compared the interventions across studies, we were unable to report evidence of publication bias. The possible effects we wanted to analyze were in relation to duration and intensity across the studies. Because outcomes from both control groups and intervention groups are considered in our study, the analytical approach does not apply to how the biases in publications are usually reported, with funnel plots of effect sizes in relation to sample size or error estimates.

3. Results

3.1. Search results

Search results are displayed in Fig. 1.

3.2. Characteristics of included studies

Tables 1 and 2 present the characteristics of the 8984 participants and the 48 studies/133 interventions included. The majority of the studies originated in the USA and other Western countries. Years of publication ranged between 1974 and 2016 with 32 of the studies published after 1999.

Thirty-three studies included participants who fulfilled criteria for alcohol use disorders by DSM. The remainder had a baseline description of a population likely to fulfill criteria for alcohol use disorder. Most studies (n = 31) recruited through contact in an alcohol outpatient setting; of these, 11 also recruited through advertisement. The remainder was recruited through a mix of inpatient settings, veteran clinics, advertisement only, and primary care. Information on previous treatment attempts, pharmacological treatment and other scores of severity of alcohol abuse was not reported consistently.

Of the 133 interventions, 53% were CBT, 10% were twelve-step facilitation, 10% were motivational interviewing, 7% were treatment as usual, while the remaining interventions included psychodynamic
therapy, educational sessions, relaxation training, short telephone calls with monitoring of alcohol use with or without feedback, non-directive reflective listening, and conjoint therapy.

Three ("Effectiveness of treatment for alcohol problems: findings of the randomised UK alcohol treatment trial (UKATT)", 2005; Lam et al., 2009; O'Farrell et al., 1992) had psychosocial functioning as a primary outcome and 33 reported it as a secondary outcome, but the reporting was very diverse and collected by non-comparable methods.

3.3. Duration and treatment effect

Actually attended weeks could not be analyzed as a predictor due to too few observations, given that weekly attendance was only reported in five studies. Attended sessions were reported in 33 studies, so there were enough observations to perform sub-analysis on this predictor.

The outcome DDD had only 38 observations and these were reported differently; thus, the number of observations where standard errors could be calculated was too low and this outcome had to be omitted from further analysis. Since none of the studies reported only DDD, no studies were excluded because of this.

The results of the meta-regression are displayed in Table 3. Neither planned duration of treatment nor frequency of planned treatment had significant associations with any of the outcome measures. Two sub-analyses were performed: 1) a meta-regression with actually attended sessions instead of planned sessions, and 2) a meta-regression with the attrition rate instead of sessions. Attrition rate was defined as the rate of actual attended sessions per planned sessions. No significant associations were found in the two sub-analyses (not shown in Table 3). Frequency of research assessments was positively and significantly associated with the two outcome measures with the highest number of observations: PDA and PHD. The coefficient indicates that increasing the frequency of research assessments with one more per month will increase PDA by between 3% and 22% (p < 0.05) and decrease PHD by between 6% and 38% (p < 0.01).

Additionally, an overview of the binary correlations between all variables is presented in Table 4.

3.4. Control variables

The binary control variables are also displayed in Table 3. The outcome measure with the most observations - percent days abstinent (PDA) and percent heavy days of drinking (PHD) was not significantly associated with any of the binary control variables. Medication in the study was significantly associated with a lower proportion of participants abstinent, as based on the 14 interventions that included medication as part of the intervention, and reporting outcome by the proportion of participants abstinent. Except for the weighted analysis by number of participants, the demographic variables were left out of the analysis to maintain a reasonably low number of model parameters.

4. Discussion

4.1. Summary of results

We have systematically reviewed and analyzed 48 studies of psychosocial interventions for alcohol use disorders (AUD) in outpatient care. The main result of the meta-regression was that planned weeks, planned sessions, actually attended sessions, attrition rate, as well as frequency of planned sessions per week, were not associated with long-term alcohol use outcomes. In addition, significant and positive associations were found between frequency of research assessments and long-term alcohol use outcomes: percent days abstinent and percent heavy days of drinking.
Table 1

Descriptive data from the 48 included studies in the systematic review and meta-analysis of duration of treatment for alcohol use disorders. Including diagnosis in population, number of participants, treatment duration and type of intervention.

<table>
<thead>
<tr>
<th>Author (ref.)</th>
<th>Year</th>
<th>Diagnostic criteria for AUD</th>
<th>Outcome measures</th>
<th>Weeks of study</th>
<th>RAS</th>
<th>NOP</th>
<th>Method of intervention</th>
<th>Planned Weeks of treatment</th>
<th>Actual Weeks</th>
<th>Planned Sessions</th>
<th>Actual sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annis and Peachey (1992)</td>
<td>1992</td>
<td>DSM-III</td>
<td>PDA, PHD, DDD</td>
<td>26</td>
<td>2</td>
<td>23</td>
<td>1</td>
<td>16</td>
<td>8</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Anton et al. (2005)</td>
<td>2005</td>
<td>DSM-IV</td>
<td>PDA, DDD</td>
<td>12</td>
<td>4</td>
<td>39</td>
<td>2</td>
<td>12</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Azrin, Sisson, Meyers, and Godley (1982)</td>
<td>1982</td>
<td>Alcoholics</td>
<td>PDA, DDD</td>
<td>26</td>
<td>7</td>
<td>14</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4.9</td>
<td>6.4</td>
</tr>
<tr>
<td>Baldin et al. (2003)</td>
<td>2003</td>
<td>DSM-IV</td>
<td>PDA, PHD, DDD</td>
<td>24</td>
<td>3</td>
<td>30</td>
<td>1</td>
<td>24</td>
<td>9</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Bowers and al-Redha (1990)</td>
<td>1990</td>
<td>Alcohol problems</td>
<td>ABS</td>
<td>26</td>
<td>3</td>
<td>8</td>
<td>1</td>
<td>9</td>
<td>8.5</td>
<td>7.4</td>
<td>7.4</td>
</tr>
<tr>
<td>Davis, Campbell, Tax, and Lieber (2002)</td>
<td>2002</td>
<td>Alc.dep.</td>
<td>ABS, DDD</td>
<td>26</td>
<td>7</td>
<td>40</td>
<td>5</td>
<td>26</td>
<td>26</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>De Wildt et al. (2002)</td>
<td>2002</td>
<td>DSM-IV</td>
<td>ABS, DDD</td>
<td>28</td>
<td>7</td>
<td>86</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>2.84</td>
<td>3.03</td>
</tr>
<tr>
<td>Dieperink et al. (2014)</td>
<td>2014</td>
<td>DSM-IV</td>
<td>PDA, ABS, PHD</td>
<td>26</td>
<td>3</td>
<td>70</td>
<td>2</td>
<td>12</td>
<td>4</td>
<td>3.2</td>
<td>3.2</td>
</tr>
<tr>
<td>Hammarberg, Wennberg, Beck, and Francck (2004)</td>
<td>2004</td>
<td>DSM-IV</td>
<td>PDA, PHD</td>
<td>24</td>
<td>4</td>
<td>32</td>
<td>5</td>
<td>24</td>
<td>4</td>
<td>2.8</td>
<td>2.8</td>
</tr>
<tr>
<td>Heathter et al. (2000)</td>
<td>2000</td>
<td>Alcohol dependence</td>
<td>PDA, ABS, DDD</td>
<td>42</td>
<td>3</td>
<td>48</td>
<td>1</td>
<td>16</td>
<td>16</td>
<td>6.56</td>
<td>6.56</td>
</tr>
<tr>
<td>Hedberg and Campbell (1974)</td>
<td>1974</td>
<td>Alcoholism</td>
<td>ABS</td>
<td>26</td>
<td>2</td>
<td>15</td>
<td>1</td>
<td>26</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Kiluk et al. (2016)</td>
<td>2016</td>
<td>DSM-IV</td>
<td>PDA, PHD</td>
<td>34</td>
<td>11</td>
<td>22</td>
<td>1</td>
<td>8</td>
<td>8</td>
<td>4.3</td>
<td>4.3</td>
</tr>
<tr>
<td>Litt, Kadden, Kabela-Cormier, and Petry (2007)</td>
<td>2007</td>
<td>DSM-IV</td>
<td>PDA, ABS</td>
<td>64</td>
<td>7</td>
<td>63</td>
<td>3</td>
<td>12</td>
<td>12</td>
<td>8.3</td>
<td>8.3</td>
</tr>
<tr>
<td>Marques and Formigoni (2001)</td>
<td>2001</td>
<td>DSM-III</td>
<td>PDA, PHD</td>
<td>68</td>
<td>2</td>
<td>77</td>
<td>1</td>
<td>34</td>
<td>17</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

(continued on next page)
<table>
<thead>
<tr>
<th>Author (ref.)</th>
<th>Year of publication</th>
<th>Diagnostic criteria for AUD</th>
<th>Outcome measures</th>
<th>Weeks of study</th>
<th>RAS</th>
<th>NOP</th>
<th>Method of intervention</th>
<th>Planned Weeks of treatment</th>
<th>Actual Weeks</th>
<th>Planned Sessions</th>
<th>Actual sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>McCrady, Epstein, and Hirsch (1999)</td>
<td>1999</td>
<td>DSM-III</td>
<td>PDA, ABS, PHD</td>
<td>48</td>
<td>8</td>
<td>30</td>
<td>1</td>
<td>15</td>
<td>15</td>
<td>22</td>
<td>15</td>
</tr>
<tr>
<td>McCrady et al. (1986)</td>
<td>1986</td>
<td>MAST &gt; 5</td>
<td>PDA, ABS, PHD</td>
<td>40</td>
<td>9</td>
<td>13</td>
<td>1</td>
<td>15</td>
<td>15</td>
<td>16.7</td>
<td>15</td>
</tr>
<tr>
<td>McKay et al. (2010)</td>
<td>2010</td>
<td>Alc.dep.</td>
<td>PDA, ABS, PHD</td>
<td>78</td>
<td>7</td>
<td>86</td>
<td>4</td>
<td>26</td>
<td>1</td>
<td>64</td>
<td>25</td>
</tr>
<tr>
<td>McKay et al. (2005)</td>
<td>2005</td>
<td>DSM-IV</td>
<td>PDA, ABS</td>
<td>104</td>
<td>7</td>
<td>102</td>
<td>5</td>
<td>16</td>
<td>16</td>
<td>135</td>
<td>16</td>
</tr>
<tr>
<td>McKay et al. (2010)</td>
<td>2010</td>
<td>Alc.dep.</td>
<td>PDA, ABS, PHD</td>
<td>78</td>
<td>7</td>
<td>86</td>
<td>4</td>
<td>26</td>
<td>1</td>
<td>64</td>
<td>25</td>
</tr>
<tr>
<td>McKay et al. (2005)</td>
<td>2005</td>
<td>DSM-IV</td>
<td>PDA, ABS</td>
<td>104</td>
<td>7</td>
<td>102</td>
<td>5</td>
<td>16</td>
<td>16</td>
<td>135</td>
<td>16</td>
</tr>
<tr>
<td>McKay et al. (2010)</td>
<td>2010</td>
<td>Alc.dep.</td>
<td>PDA, ABS, PHD</td>
<td>78</td>
<td>7</td>
<td>86</td>
<td>4</td>
<td>26</td>
<td>1</td>
<td>64</td>
<td>25</td>
</tr>
<tr>
<td>McKay et al. (2005)</td>
<td>2005</td>
<td>DSM-IV</td>
<td>PDA, ABS</td>
<td>104</td>
<td>7</td>
<td>102</td>
<td>5</td>
<td>16</td>
<td>16</td>
<td>135</td>
<td>16</td>
</tr>
<tr>
<td>McKay et al. (2010)</td>
<td>2010</td>
<td>Alc.dep.</td>
<td>PDA, ABS, PHD</td>
<td>78</td>
<td>7</td>
<td>86</td>
<td>4</td>
<td>26</td>
<td>1</td>
<td>64</td>
<td>25</td>
</tr>
<tr>
<td>McKay et al. (2005)</td>
<td>2005</td>
<td>DSM-IV</td>
<td>PDA, ABS</td>
<td>104</td>
<td>7</td>
<td>102</td>
<td>5</td>
<td>16</td>
<td>16</td>
<td>135</td>
<td>16</td>
</tr>
<tr>
<td>McKay et al. (2010)</td>
<td>2010</td>
<td>Alc.dep.</td>
<td>PDA, ABS, PHD</td>
<td>78</td>
<td>7</td>
<td>86</td>
<td>4</td>
<td>26</td>
<td>1</td>
<td>64</td>
<td>25</td>
</tr>
</tbody>
</table>

### Notes
- **YOP**: Year of publication.
- **AUD**: Diagnostic and statistical manual. **MAST**: Michigan alcohol screening test.
- **PDA**: Percentage of days abstinent. **PHD**: Percentage of heavy days drinking (Heavy drinking was defined as 4 or more drinks per day for women and 5 or more drinks per day for men in most of the studies. A few had a definition in oz. (≥3 oz.) or grams (≥68–80 g.) of alcohol and one reported it as a day where the participant felt intoxicated). **ABS**: Share of participants abstinent (definition varies between one month abstinent and up to one year either abstinent or with few light drinking days). **DDD**: Drinks per drinking day.
- **RAS**: Research assessments.
- **NOP**: Number of participants.
4.2. One size does not fit all

Overall, our findings are in line with earlier reviews of planned duration and effect of treatment of substance use disorders (Blodgett et al., 2014; Lenaerts et al., 2014; McKay, 2005, 2009).

One explanation of the lack of significant findings may be the heterogeneity of people suffering from AUD. Such heterogeneity coheres with the good-enough level model, whereby people recover and their symptoms diminish at different rates (Barkham et al., 2006). This would suggest that the dose-effect model is ill-fitted for the analysis of the effectiveness of duration of treatment among individuals with AUD.

Comparisons of the two models within psychotherapy in general indicate that the dose-effect model fits well at the beginning of therapy (Baldwin, Berkeljon, Atkins, Olsen, & Nielsen, 2009; Delgadillo et al., 2014), but that the good-enough level model is a better fit with longer treatment durations beyond six (Delgadillo et al., 2014) to eight (Baldwin et al., 2009) sessions. On the other hand, these trajectories of change seem interchangeable and more diverse than the models may be able to grasp (Owen et al., 2015). Furthermore, it is likely that patients are influenced by information about the planned duration of therapy and may partly adjust their trajectory of change accordingly. A constraint on research into the good-enough level model is that it...
requires reporting of outcome at the individual level: session to session, or at a minimum, by the end of actually attended treatment.

The difficulty of identifying a successful treatment approach for the patient population as a whole has long led to the conclusion that “one size does not fit all” (Institute of Medicine Committee on Treatment of Alcohol, 1990). Instead, the hypothesis has been put forward that some treatments are possibly more attractive and effective in targeting certain specfic patient groups than others. It has also been postulated that the variance in the duration of treatment is too small for significant associations between duration and effect to be found— which constitutes an argument for a larger timescale in studies of duration in therapy (Buhringer, 2006). Moreover, it is important to note that people who drop out of treatment are often counted as relapse cases; however, some dropouts might have rapid trajectories of change and may thus be considered as successes.

We argue that the complexity of AUD requires both subgrouping of the population and the measurement of improvements session to session— thereby also reporting the actually attended sessions—in order to study and optimize the appropriate treatment length. If there are associations between duration and effect of treatment within subgroups, the next question would be: “What, then, is the optimal treatment duration for this group, and how well does this match with current recommendations in guidelines?” An optimal treatment duration is necessary as an argument in the budgetary planning and logistics of treatment centers. If the optimal treatment duration is unknown, flexibility is called for— a flexibility treatment centers may not always be able to provide, with burdensome consequences at psychosocial and societal levels (Cuijpers et al., 2013).

Another aspect concerns how the effect of treatment is assessed. As discussed by Pearson et al. (Pearson, Kirouac, & Witkiewitz, 2016), the quantity of alcohol use is not part of the DSM criteria, yet most studies primarily focus on the quantity of alcohol as a measure of effect of treatment. The DSM criteria (Association, 2013) are all related to drinking as such— the quantity is not defined; hence, only a treatment outcome of full abstinence would be equivalent to no DSM symptoms. As studied by Wilson et al. (2016), when measuring effect of treatment at different levels of alcohol consumption, there is a risk of too many false positives and false negatives in levels of psychosocial functioning. Moreover, studies of the association between drinking levels and consequences from drinking do not reveal a logical cut-off point in drinking levels (Pearson, Bravo, Kirouac, & Witkiewitz, 2017). Since many treatment centers accept controlled drinking as a treatment goal, adding an outcome reflecting psychosocial functioning may be appropriate (Wilson et al., 2016). A binary classification of problematic drinking does not take into account the level of suffering, the level of blood alcohol content in the individual, nor the risk of harm— again, one size does not fit all (Pearson et al., 2016).

### 4.3. Research assessments and effect

Our findings support the recommendation for a low frequency of research assessment in studies of AUD (Clifford et al., 2007). Research assessments are a hardly avoidable bias in controlled studies and it may be that research assessments have reduced the ability to detect significant differences between groups in some cases. For example, there has been speculation about the impact of research assessments in relation to the Project MATCH (Buhringer, 2006).

Research assessments may function like control visits similar to Project MATCH (Buhringer, 2006). In that respect, research assessments are a hardly avoidable bias in controlled studies and it may be that research assessments have reduced the ability to detect significant differences between groups in some cases. For example, there has been speculation about the impact of research assessments in relation to the Project MATCH (Buhringer, 2006).
4.4. Limitations

There are limitations that may mask possible associations between duration and effect of treatment in long-term alcohol use:

Despite well-defined in- and exclusion criteria, the heterogeneity of study features and population was substantial. By assuming treatment uniformity, the data does not apply to a particular type of intervention. Additionally, we have chosen not to differentiate between methods in the studies regarding recruitment, quality control of the therapy, and the blinding of participants and researchers. We did not have a control group and were, therefore, not able to control for publication bias. To take the heterogeneity of the population into account, some factors reported consistently in the studies were included in the analysis, but other factors, especially severity of AUD, could not be included.

Because of heterogeneity in the reporting of psychosocial functioning, it was not possible to perform sub-analysis of the associations between psychosocial functioning and duration of treatment. We were not able to account for the actually attended weeks, and the sub-analysis of actually attended sessions and attrition rate left out one third of the interventions.

The time between screening and baseline as part of the research study duration could also be important, but this was rarely reported and thus further compromises the accuracy of the data on duration. Moreover, the duration of the research study was in some cases longer, with a research assessment point at 52 weeks as one example, but with primary outcome at 26 weeks. This last research assessment was not counted in the study in several aspects (actually attended weeks and sessions, and attrition rate left out one third of the interventions).

4.5. Strengths

Our analysis is based on a substantial sample size of people suffering from alcohol use disorders with mean descriptive factors comparable to the populations treated in many real-life alcoholic outpatient care centers. As many variables as possible were taken into account in the meta-regression, showing diverse effects related to subgroups. This study visualizes the complexity of recommending a specific duration of psychosocial treatment for such a diverse group as people with AUD.

5. Perspective

Our review raises methodological considerations for future studies, notably the need for more consistent and detailed reporting of the characteristics of the participants, rate of refusal to participate, duration of the study in several aspects (actually attended weeks and sessions, and duration from first approach to participants until baseline), as well as description and quality assessment of all interventions including the control condition. We recommend that data be collected at an individual level and session by session to enable analysis for optimal individualized treatment duration and effect of treatment. If not reported, we recommend publicly sharing data at the individual level to facilitate secondary analysis. For a more multifaceted aspect of treatment effects, there is a need for consensus on how to report psychosocial functioning and to include psychosocial functioning as an outcome presented with
effect sizes. Future research should target treatment of subgroups of patients with alcohol use disorder, thereby reducing the heterogeneity of groups compared. Low frequency of research assessments is encouraged if we wish to assess the impact of the treatment as such. Finally, randomized controlled studies of different durations of the same intervention would be ideal.

Financial support
This study is funded unconditionally by the Lundbeck Foundation, the Region of Southern Denmark, and the University of Southern Denmark.

References


