Title

A Narrative Review of Outcome Studies for Residential and Partial Hospital-based Treatment of Eating Disorders.

Journal

European eating disorders review : the journal of the Eating Disorders Association, 24(4)

ISSN

1072-4133

Authors

Friedman, Keren
Ramirez, Ana L
Murray, Stuart B
et al.

Publication Date

2016-07-01

DOI

10.1002/erv.2449

Peer reviewed

REVIEW

A Narrative Review of Outcome Studies for Residential and Partial Hospital-based Treatment of Eating Disorders

Keren Friedman[1], Ana L. Ramirez[2], Stuart B. Murray[2,3], Leslie K. Anderson[2], Anne Cusack[2], Kerri N. Boutelle[2] & Walter H. Kaye[2]*

[1]Sackler School of Medicine, Tel Aviv University, Israel
[2]Department of Psychiatry, University of California, San Diego, CA, USA
[3]Department of Psychiatry, University of California San Francisco, San Francisco, CA, USA

Abstract

Objective: The objective of this study was to review the current eating disorders outcome literature after residential or partial hospitalization programme (PHP) treatment.
Method: Articles were identified through a systematic search of PubMed and PsycINFO.
Results: Twenty-two PHP and six residential treatment studies reported response at discharge and tended to find improvement. Fewer studies (nine PHP and three residential) reported outcome at some interval after discharge from treatment. These tended to find sustained improvement. A substantial proportion of patients were lost to follow-up, particularly for residential treatment. Only two follow-up studies used controlled trials; both showed efficacy for PHP compared with inpatient treatment with regard to maintaining symptom remission.
Conclusions: Improvement at discharge may not predict long-term outcome. Long-term follow-up studies were confounded by high dropout rates. While higher levels of care may be essential for reversing malnutrition, there remains a lack of controlled trials showing long-term efficacy, particularly for residential treatment settings.

Accepted 5 March 2016; Received 12 November 2015; Revised 23 January 2016

Keywords
anorexia nervosa; bulimia nervosa; outcome; treatment

***Correspondence**
Walter H. Kaye, MD, UCSD Eating Disorders Center for Treatment and Research, Department of Psychiatry, University of California, 4510 Executive Drive, Suite 315, San Diego, CA 92121, USA. Tel: 858 534 3951; Fax: 858 534 6070.
Email: wkaye@uscd.edu

Published online in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/erv.2449

Eating disorders (ED) such as anorexia nervosa (AN), bulimia nervosa (BN) and eating disorder not otherwise specified (EDNOS) continue to rank amongst the most complex and pernicious of all psychiatric illnesses, posing significant challenges for both the patient and their family or caregivers. ED often exhibit poor treatment outcomes (Steinhausen, 2009), high rates of relapse (Grilo et al., 2012), high rates of treatment dropout (Dejong, Broadbent, & Schmidt, 2012) and elevated rates of comorbid medical complexities (Berkman, Lohr, & Bulik, 2007).

Clinical practice has witnessed a shift towards a more individualized management of ED via stepped levels of clinical care (Wilson, Vitousek, & Loeb, 2000). Recently, suggestions have been made for ED treatment dosage to be determined by illness severity (Maguire et al., 2008). A multitude of treatment platforms may best meet the needs of differing degrees of illness. While offering the minimal level of clinical care required for recovery, this also reduces the cost of treatment (Kaye, Enright, & Lesser, 1998; Wiseman, Sunday, Klapper, Harris, & Halmi, 2001). A patient's transition through levels of care is typically determined by factors such as symptom severity, medical status, motivational status, treatment history and financial influences (Kaplan, Olmsted, Carter, & Woodside, 2001). Movement through levels of care may be bidirectional, depending on the patient's needs. The currently established continuum of clinical care for those with ED includes outpatient, intensive outpatient, partial hospitalization programme (PHP), residential and inpatient treatment settings (Murray et al., 2015), with PHP and residential settings having recently emerged as viable clinical options to bridge the gap between inpatient and traditional outpatient services.

Residential treatment programmes house patients living full time in non-hospital-based treatment settings, where they typically receive meal support, multidisciplinary treatment and individual and group therapy, with treatment dose averaging approximately 83 days (Frisch, Franko, & Herzog, 2006). This treatment differs from inpatient treatment, which could imply admission into a medical unit at a hospital or into a psychiatric hospital. Alternately, PHP settings are characterized by patients spending 6–10 hours a day, between 3 and 7 days a week, in an outpatient programme where they receive daily meals, snacks and group therapy, as well as regular individual therapy and

dietetic and medication management (Abbate-Daga et al., 2009).

Partial hospitalization programme and residential programmes have recently become more prevalent (Twohig, Bluett, Torgesen, Lensegrav-Benson, & Quakenbush-Roberts, 2014). However, despite the increased presence of these facilities in our clinical landscape, little is known about their long-term effectiveness. To date, only two reviews have been published, and both have limitations. One review (Zipfel et al., 2001) was published in 2001 and reported on 39 published studies exploring outcomes of PHPs. Whereas this review demonstrates preliminary support for the treatment of ED in PHP settings, it also illustrates the absence of a large body of empirical evidence. A more recent review of adult PHP treatment notes that upon discharge from the programmes, adults with ED experienced an increase in body mass index (BMI), a reduction in psychological symptoms and an increase in self-esteem (Hepburn & Wilson, 2014). However, this review includes adult studies only and reports outcomes at time of discharge only, without follow-up assessments at some point in time after discharge. Thus, the purpose of the present paper is to provide an up-to-date review of the efficacy of both residential and PHP treatment settings in the treatment of a range of ED with both adolescent and adult presentations, integrating follow-up data wherever possible.

Methods

Search, study selection and inclusion criteria

A systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009). The electronic databases PubMed and PsycINFO were searched to identify studies examining the outcome of PHP and residential treatment programmes for adolescents and adults with ED that were published subsequent to the most recent comprehensive review in 2001 (Zipfel et al., 2001). The date of our last literature search was 1 July 2015. Both databases were searched by typing the following combinations of words into the search bar: 'anorexia' or 'bulimia' or 'eating disorder' with 'day hospital', 'day treatment', 'partial hospitalization' or 'residential'. Studies were screened first by title, second by abstract and third according to content. Screening was performed according to the inclusion criteria outlined in Table SIV in the Supporting Information. As most studies in the review were completed before the publication of the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders*, we included studies that utilized the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition, text revision. Notwithstanding the difficulties brought about by the widely varying measures of outcome between studies, 22 PHP studies and 7 residential studies were determined to be eligible for this review (Figure 1). Zeeck

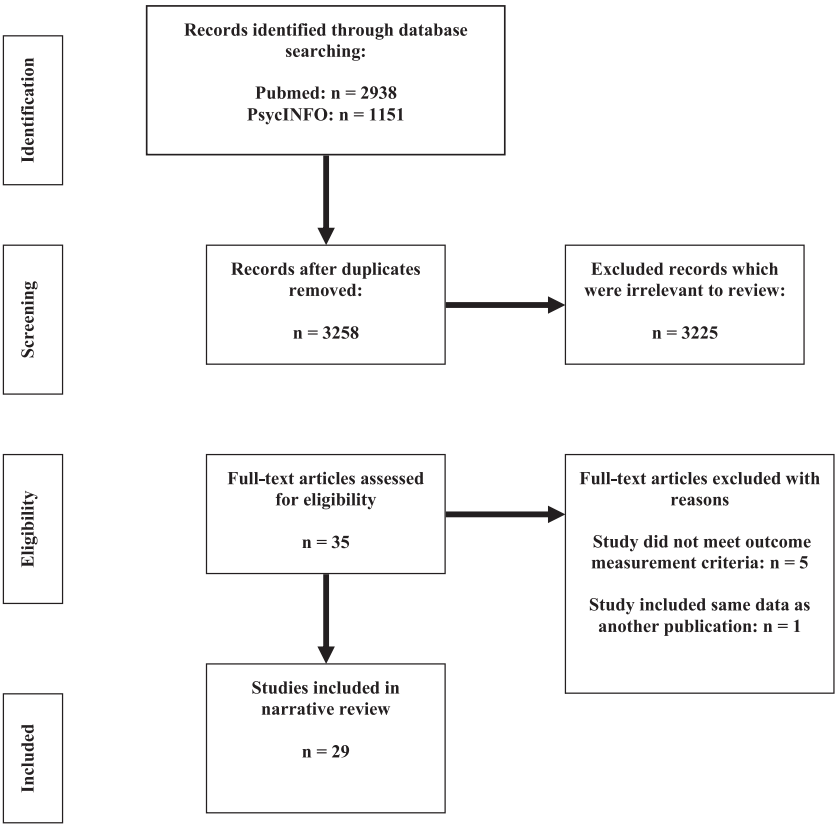

Figure 1. PRISMA flowchart illustrating the identification and screening of relevant studies

et al. published two studies with the same patient sample (Zeeck, Weber, Sandholz, Joos, & Hartmann, 2011; Zeeck et al., 2009). We therefore include the 2011 study in the review, although we also collected details from the 2009 paper. When necessary, the study authors were contacted to obtain additional information that was not provided in the published article (Lowe, Davis, Annuziato, & Lucks, 2003). Where indicated, we report summary data in terms of differences in means at intake, discharge and follow-up. However, it is important to note that because of the self-report nature of many measures of ED pathology and the inherent biases within studies (Murray, Loeb, & Le Grange, 2014), selective reporting bias could not be entirely excluded.

Data collected from all 29 studies are presented in Tables 1–3. The ages, weights or BMIs, questionnaire scores and binging and purging frequencies are presented as means, with standard deviations (when provided) in parentheses. Significance levels are indicated by symbols, described in each table.

Assessment of quality and risk of bias

We assessed the risk of bias within individual studies and performed a quality appraisal by using the Effective Public Health Practice Project (EPHPP) quality assessment tool, which uses seven categories to evaluate study quality (National Collaborating Centre for Methods and Tools, 2008) (see Table SV in the Supporting Information). These seven scales include selection bias, study design, confounders, blinding, data collection methods, withdrawal and dropouts, intervention integrity and analyses. Based on the scores, each study receives a 'global rating'. The scales 'confounders' and 'intervention integrity' were omitted, as they were difficult to assess for all of the studies reviewed and would not add substantial data. For 'analyses', studies were given a rating of 'moderate' if appropriate statistical analyses were performed. They were rated 'strong' if they included intention to treat analyses. Lastly, for each study that included follow-up data, we gave two global ratings—one for the discharge data and one for the follow-up data. Two authors (K. F. and A. L. R.) independently rated each study in order to improve the quality of our assessment.

Results

Analysis of study quality assessment

The results of the quality assessment conducted by the EPHPP tool are presented in Table SV in the Supporting Information. Both authors (K. F. and A. L. R.) agreed on the majority of scales. Disagreements were discussed, and revisions were made accordingly. The global ratings for discharge studies/data were mostly 'moderate' with the exception of three studies, which were rated globally as 'weak'. In terms of follow-up studies/data, half received a global score of 'moderate', while the other half received a global score of 'weak'.

Programme characteristics

Partial hospitalization programmes

Although PHPs share the goals of normalization of weight, normalization of eating behaviour and identification and resolution of ED behaviours, PHP protocols vary widely amongst different treatment centres and include distinct and diverse therapeutic approaches (Willinge, Thornton, Olmsted, & Touyz, 2012). Of the PHP studies we reviewed, two (Goddard et al., 2013; Herpertz-Dahlmann et al., 2014) were multicentre studies and did not provide information regarding their clinical content. The remaining PHP studies did report details regarding their respective treatment modalities (Abbate-Daga et al., 2015; Crino & Djokvucic, 2010; Dancyger et al., 2003; Exterkate, Vriesendorp, & de Jong, 2009; Fittig, Jacobi, Backmund, Gerlinghoff, & Wittchen, 2008; Girz, LaFrance Robinson, Foroughe, Jasper, & Boachie, 2013; Goldstein et al., 2011; Henderson et al., 2014; Hoste, 2015; Jones, Bamford, Ford, & Schreiber-Kounine, 2007; Kong, 2005; Manara, Manara, & Todisco, 2005; Olmsted, Kaplan, & Rockert, 2003; Olmsted, McFarlane, Trottier, & Rockert, 2013; Ornstein, Lane-Loney, & Hollenbeak, 2012; Schaffner & Buchanan, 2008; Treat, McCabe, Gaskill, & Marcus, 2008; Willinge, Touyz, & Thornton, 2010; Zeeck, Herzog, & Hartmann, 2004; Zeeck et al., 2011). All of these programmes included group therapy sessions, and some included group meals (Olmsted et al., 2003; Olmsted et al., 2013; Schaffner & Buchanan, 2008), skills groups (Fittig et al., 2008; Girz et al., 2013; Goldstein et al., 2011; Kong, 2005; Schaffner & Buchanan, 2008) and relaxation groups (Jones et al., 2007; Schaffner & Buchanan, 2008; Zeeck et al., 2011). In addition, 14 studies included cognitive behaviour therapy (CBT) and/or dialectical behaviour therapy groups (Crino & Djokvucic, 2010; Exterkate et al., 2009; Fittig et al., 2008; Goldstein et al., 2011; Herpertz-Dahlmann et al., 2014; Jones et al., 2007; Kong, 2005; Manara et al., 2005; Olmsted et al., 2003; Olmsted et al., 2013; Treat et al., 2008; Willinge et al., 2010; Zeeck et al., 2004; Zeeck et al., 2011). Three studies were identified as primarily psychodynamic (Abbate-Daga et al., 2015; Zeeck et al., 2004; Zeeck et al., 2011). Dietician consultations and meal planning were frequently incorporated as part of the treatment plan in PHP settings. Twelve of these studies specifically described such components (Abbate-Daga et al., 2015; Girz et al., 2013; Goldstein et al., 2011; Henderson et al., 2014; Jones et al., 2007; Olmsted et al., 2003; Olmsted et al., 2013; Schaffner & Buchanan, 2008; Treat et al., 2008; Willinge et al., 2010; Zeeck et al., 2004; Zeeck et al., 2011). Lastly, family-based treatment (FBT) was a component of several PHPs, particularly, for adolescent patients (Girz et al., 2013; Henderson et al., 2014; Hoste, 2015; Ornstein et al., 2012). Other programmes, while not specifically implementing manualized FBT, did include family participation as a part of the treatment plan (Fittig et al., 2008; Goldstein et al., 2011; Kong, 2005; Schaffner & Buchanan, 2008; Treat et al., 2008; Willinge et al., 2010; Zeeck et al., 2004; Zeeck et al., 2011).

In general, PHPs operate between 6 and 12 hours a day, for 4–7 days a week, although little evidence has demonstrated the optimal intensity of treatment dosage. One study reported that a 4-day and a 5-day programme were equally effective in bringing about weight gain, although it noted that the 5-day programme resulted in a greater reduction in binge/purge episodes and further improved the quality of life (Olmsted et al., 2013). A follow-up study a decade later corroborated these earlier findings, although it concluded that the optimal treatment intensity varied according to the patient's goals (Olmsted et al., 2013). Moreover, little consensus exists regarding the optimal duration of treatment in PHP settings. Of the studies reviewed, treatment duration

Table 1 Summary of discharge results for open trials

Author	PHP or residential	Initial sample N	Discharge data (N)	Discharge sample (%)	Age (mean and *SD*)	Diagnosis	Primary therapy	Average treatment duration (days)	Baseline data (weight and/or behaviour)	Outcome (weight)	Outcome (behaviour)
Abbate-Daga et al. (2015)	PHP	56	50	89	25 (6) years	AN	Psychodynamic	180	BMI = 16.3 (2.7) EDI (total) = 107.0 (55.1)	BMI = 17.3 (2.7)**	EDI (total) = 89.1 (51.1)***
Hoste (2015)	PHP	28	21	75	17 (4) years	AN	FBT	32	%EBW = 82.1 (9.6) EDE-Q (global) = 3.2 (1.9)	EBW = 93.1%	EDE-Q (global) = 1.9 (1.4)**
Henderson et al. (2014)	PHP	65	65	100	15 (1) years	AN, BN, EDNOS	FBT	104	BMI (all ED groups) = 18.7 (2.4) EDI-2 (DT) = 16.1 (6.0)	BMI = 20.5 (2.0)***	EDI-2 (DT) = 11.6 (7.4)***
Goddard et al. (2013)	PHP	16	15	94	25 (5) years	AN	Varied among centres	Varied among centres	BMI = 17.3 (2.4) EDE-Q = 4.4 (1.0)	BMI = 18.5 (2.2)	EDE-Q = 3.7 (1.3)**
Olmsted et al. (2013)	PHP	1294	1074	83	26 (8) years	AN BN, EDNOS	CBT	74	BMI (AN) = 16.9 Binging (BN) = 46.4 Purging (BN) = 72.4 EDI-2 (DT) = 14.9	BMI (AN) = 19.3***	Binging (BN) = 4.0** Purging (BN) = 5.7** EDI-2 (DT) = 9.9***
Girz et al. (2013)	PHP	17	17	100	16 (1) years	AN, BN, EDNOS	FBT	150	IBW = 88.0% EDI-2 (DT) = 49.2 (12.6)	IBW = 16/17 met IBW, the rest = 99.0%	EDI-2 (DT) = 31.1 (13.1)**
Ornstein et al. (2012)	PHP	30	30	100	13 (2) years	AN, BN, EDNOS	FBT	72	BMI (all ED groups) = 16.5 (2.3) ChEAT (total) = 20	BMI (all ED groups) = 18.4 (1.6)***	ChEAT (total) = 9.0***
Goldstein et al. (2011)	PHP	28	26	93	15 (not reported) years	AN, EDNOS	CBT	70	BMI = 16.5 (1.5) EDI-2 (DT) = 13.8 (9.1)	BMI = 17.1 (1.9)*	EDI-2 (DT) = 10.1 (8.3)***
Crino and Djokvucic (2010)	PHP	36	33	92	26 (9) years	AN BN, EDNOS	CBT	84	BMI (underweight group) = 16.6 (1.5) Binging (B/P group) = 50.4 (47.5) Purging (B/P group) = 77.8 (76.0) EAT = 34.3 (12.6)	BMI (underweight group) = 17.7 (1.5)***	Binging (B/P group) = 13.0 (21.1)** Purging (B/P group) = 23.8 (30.5)*** EAT = 22.7 (10.1)***
Willinge et al. (2010)	PHP	58	44	76	22 (5) years	AN BN, EDNOS	CBT	33	BMI (underweight group) = 17.6 (1.9) Binging (B/P group) = 21.2 (33.9) Purging (B/P group) = 9.1 (16.7) EDI-2 (DT) = 21.0 (6.6)	BMI (underweight group) = 19.4 (2.0)***	Binging (B/P group) = 3.0 (5.3)*** Purging (B/P group) = 0.7 (1.4)** EDI-2 (DT) =13.3 (7.5)**
Exterkate et al. (2009)	PHP	319	193	61	26 (6) years	AN, BN, EDNOS	CBT	182	BMI (AN) = 15.8 EDES (all ED groups) = 39.0	BMI (underweight group) = 19.2***	EDES (all ED groups) = 55.7***
Fittig et al. (2008)	PHP	283	199	70	22 (4) years	AN, BN	CBT	140	BMI (AN) = 15.8 Binging (BN) = 2.2 Purging (BN) = 2.0 EDI-2 (DT) (all ED groups) = 12.0	BMI (all AN) = 18.9***	Binging (BN) = 0.5*** Purging (BN) = 0.4*** EDI-2 (DT) (all ED groups) = 3.3***

Continues

Table 1. (Continued)

Schaffner and Buchanan (2008)	PHP	77	66	86	21 (7) years	Not specified	CBT	90	Weight = 117.1 (33.3) Binging (BN) = 7.1 (10.3) Purging (BN) = 10.1 (12.1) EDI-2 = 8.1 (3.8)	Weight = 124.5 (30.1)***	Binging (BN) = 0.3 (0.5)** Purging (BN) = 0.5 (1.7)*** EDI-2 = 3.1 (2.8)***
Treat et al. (2008)	PHP	71	71	100	18 (6) years	AN	CBT	22	BMI = 15.2 (1.5) EDI-2 (DT) = 13.2 (6.9)	BMI = 17.9 (1.1)**	EDI-2 unreported
Jones et al. (2007)	PHP	34	22	65	26 (6) years	AN BN, EDNOS	CBT	84	BMI (AN) = 16.4	Average increase in BMI = +1.6 (1.6)***	Average EDE = Q improvement = 1.2 (1.5)***
Manara et al. (2005)	PHP	46	41	89	23 (6) years	AN, BN	CBT	97	BMI (AN) = 15.7 (1.7) EDI-2 (DT) (AN) = 13.6 (7.6) EDI-2 (bulimia) (BN) = 11.3 (6.0)	BMI (AN) = 18.5(1.4)***	EDI-2 (DT)(AN) = 3.6 (5.6)*** EDI-2 (bulimia) (BN) = 0.0 (0.0)***
Zeeck et al. (2004)	PHP	19	18	95	27 (7) years	BN	Psychodynamic	86	EDI-2 (bulimia) = 13.3 (4.4)	N/A	EDI-2 (bulimia) = 6.2 (5.9)**
Dancyger et al. (2003)	PHP	82	40	49	18 (3) years	AN, BN, EDNOS	Multidisciplinary	107	IBW% (AN) = 87.0% EDI-2 (DT) = 14.7	IBW% (AN) = 91.0%	EDI-2 unreported
Olmsted et al. (2003)	PHP	756	639	85	26 (7) years	AN BN, EDNOS	CBT	74	BMI (BMI < 18.5) = 16.8 Binging (BN) = 48.2 Purging (BN) =77.4 EDI-2 (DT) = 15.1	BMI (for BMI < 18.5) = 19.3**	Binging (BN) = 3.7*** Purging (BN) = 5.5 EDI-2 (DT) = 9.5***
Weltzin et al. (2014)	Residential	148	105	71	15 (2) years	AN, BN	CBT	56	BMI = 18.7 (3.5) EDE-Q (global) = 3.7 (1.6)	BMI = 20.6 (2.9)	EDE-Q (global) = 2.0 (1.4)***
Brewerton and Costin (2011b)	Residential	215	215	100	31 (8) years	AN, BN	CBT	86	BMI (AN) = 15.9 (1.8) EDI-2 (DT) (AN) = 13.9 Binging (BN) = 4.5 Purging (BN) = 4.2	BMI (AN) = 18.2 (1.4)***	EDI-2 (DT) (AN) = 8.5* Binging (BN) = 0.6** Purging (BN) = 0.3**
Delinsky et al. (2010)	Residential	68	36	53	19 (2) years	AN, BN	CBT	50	EBW = 80.6% EDE-Q (global) = 4.0 (1.3) Binging = 10.4 (3.6) Purging = 8.4 (11.3)	EBW = 91.8%***	EDE-Q (global) = 2.6 (1.4)*** Binging = 3.6 (4.3) Purging = 1.3 (2.8)***
Hoffart et al. (2010)	Residential	39	36	92	29 (7) years	BN	CBT	105	EDE-Q used for baseline binging and purging, numbers not reported	N/A	Binging = $z = -2.6$* Purging = $z = 1.8$
McHugh (2007)	Residential	65	65	100	17 (2) years	AN	Not reported	29	BMI = 15.8 (1.3) EDI-2 (DT) = 12.9 (6.5)	BMI = 17.8 (1.2)***	EDI-2 (DT) = 9.4 (6.9)***
Lowe et al. (2003)	Residential	917	470	51	25 (8) years	AN, BN	Psychodynamic, CBT	19	BMI (AN) = 16.2 EAT-D (AN) = 23.9 (9.7) EAT-B (BN) = 12.0 (4.5)	BMI (AN) = 17.9***	EAT-D (AN) = 17.1 (10.5) *** EAT-B (BN) = 6.2 (5.3)***
PHP											
Total		3315	2664								
Average		174	140	84	21			93			
SD		324	268	15	5			46			
Residential											
Total		1452	927								
Average		242	155	78	23			58			
SD		337	168	23	7			33			

Note: BMI, body mass index; EAT, Eating Attitudes Test; EDI, Eating Disorder Inventory (DT, drive for thinness subscale); EDE-Q, Eating Disorder Examination; EDES, Eating Disorder Evaluation Scale; EBW, expected body weight; IBW, ideal body weight; ChEAT, The Children's Eating Attitudes Test; N/A, not applicable. This table describes the discharge data for the PHP and residential open trials in this review. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

Table 2 Summary of follow-up data for open trials

Author	Type (PHP or residential)	Initial sample N	No. followed up			Total FU (N)	% FU	Time to FU (months)	Baseline data (weight and/or behaviour)	Outcome (weight)	Outcome (behaviour)
			BN	AN	EDNOS						
Abbate-Daga et al. (2015)	PHP	56		48		48	86	12	BMI = 16.3 (2.7) EDI-2 (total) = 107.0 (55.1)	BMI = 17.3 (2.7)**	EDI-2 (total) = 85.0 (59.8)**
Henderson et al. (2014)	PHP	65	61			61	94	6	BMI (all ED groups) = 18.7 (2.4) EDI-2 (DT) = 16.1 (6.0)	BMI = 19.8 (2.2)***	EDI-2 (DT) = 11.7 (7.3)***
Goldstein et al. (2011)	PHP	28		20		20	71	6	BMI = 16.5 (1.5) EDI-2 (DT) = 13.8 (9.1)	BMI: 18.3 (2.5)**	EDI-2 (DT) = 5.9 (6.9)***
Willinge et al. (2010)	PHP	58	27			27	47	3	BMI (underweight group) = 17.6 (1.9) Binging (B/P group) = 21.2 (33.9) EDI-2 (DT) = 13.0 (9.0) *** Purging = 9.1 (16.7) EDI-2 (DT) = 21.0 (6.6)	BMI (underweight group) = 19.5 (2.6)***	Binging (B/P group) = 3.8 (11.5)** Purging (B/P group) = 3.4 (11.5)**
Fittig et al. (2008)	PHP	283	57	64		121	43	18	BMI (AN) = 15.8 Binging (BN) = 2.2 Purging (BN) = 2.0 EDI-2 (DT) (all ED groups) = 12.0	BMI (AN) = 19.4***	Binging (BN) = 0.5*** Purging (BN) = 0.4*** EDI-2 (DT) (all ED groups) = 3.0***
Treat et al. (2008)	PHP	71		37		37	52	6	BMI = 15.2 (1.5) EDI-2 (DT) = 13.2 (6.9)	N/A	47.9% of patients maintained outpatient treatment status
Jones et al. (2007)	PHP	34	15			15	44	3	BMI (AN) = 16.4	50% showed improvement in weight; 20% were at a healthy weight	12/15 showed continued improvement (EDE-Q)
Manara et al. (2005)	PHP	46	17	20		37	80	6	BMI (AN) = 15.7 (1.7) EDI-2 (DT)(AN) = 13.6 (7.6) EDI-2 (bulimia) (BN) = 11.3 (6.0)	BMI (AN) = 18.1 (1.8)***	EDI-2 (DT)(AN) = 5.0 (5.8) *** EDI-2 (bulimia) (BN) = 2.0 (5.0)***
Zeeck et al. (2004)	PHP	18	14			14	78	18	EDI-2 (bulimia) = 13.3 (4.4)	N/A	Numbers not reported; 50% in remission
Weltzin et al. (2014)	Residential	148	61			61	41	24	BMI = 18.7 (3.5) EDE-Q (global) = 3.7 (1.6)	BMI = 21.0	EDE-Q (global) = 2.0 (1.7)***
Brewerton and Costin (2011a)	Residential	215	52	66		108	55	50	BMI (AN) = 15.9 (1.8) EDI-2 (DT) (AN) = 13.9 Binging (BN) = 4.5 Purging (BN) = 4.2	BMI (AN) = 19.0 (3.3)***	EDI-2 (DT) (AN) = 7.0* Binging (BN) = 1.7 Purging (BN) = 1.3**
Lowe et al. (2003)	Residential	917	52	98		150	16	3	BMI (AN) = 16.2 EAT-D (AN) = 23.9 (9.7) EAT-B (BN) = 12.0 (4.5)	BMI (AN) = 18.0***	EAT-D (AN) = 17.2 (10.9)*** EAT-B (BN) = 6.6 (4.9)**
PHP											
Total		659				380					
Average		73				42	66	9			
SD		81				33	20	6			
Residential											
Total		1280				319					
Average		427				106	37	26			
SD		426				45	20	24			

Note: BMI, body mass index; EAT, Eating Attitudes Test; EDI, Eating Disorder Inventory (DT, drive for thinness subscale); EDE-Q, Eating Disorder Examination; EDES, Eating Disorder Evaluation Scale; EBW, expected body weight; IBW, ideal body weight; FU, follow-up; N/A, not applicable. This table describes the follow-up data for the PHP and residential open trials in this review. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

Table 3 Summary of PHP RCTs

Author	Control group treatment	Initial sample in PHP (N)	Discharge data (N) for PHP	Discharge sample (%)	Age (mean and *SD*) for PHP treatment	Diagnosis	Primary therapy	Baseline data (weight and/or behaviour) for PHP	Discharge outcome data (weight and/or behaviour) for PHP	FU data N (%) for PHP	Time to FU (months)	FU outcome data (weight and/or behaviour)
Herpertz-Dahlmann et al. (2014)	Inpatient	87	87	100	15 (2) years	AN	CBT	BMI = 14.9 (1.5) EDI-2 (global) = 248.8 (58.2)	BMI = 18.1 (0.9)***	86 (99%)	12	BMI = 18.1 (2.0) *** EDI-2 (global) = 248.2 (71.1) BMI in PHP was not inferior to BMI outcome in inpatient treatment
Zeeck et al. (2011)	Inpatient	22	22	100	26 (8) years	BN	Psychodynamic	EDI-2 (bulimia) = 11.6 (3.7) Binging severity = 2.5 (0.8) Purging severity = 2.9 (0.4)	EDI-2 (bulimia) = 5.1 (3.9)*** Binging severity = 2.5 (0.8)*** Purging severity = 2.9 (0.4)***	21 (95%)	36	EDI-2 (bulimia) = 3.8 (4.9)*** 72% of PHP patients in remission, 69% of inpatient patients in remission
Kong (2005)	Outpatient	25	21	84	27 (7) years	AN, BN, EDNOS	CBT	BMI = 21.5 (5.7) Binging = 9.8 (4.0) Purging = 6.1 (5.8) EDI (DT) = 12.4 (4.7) EDI (bulimia) = 12.3 (5.9)	BMI = 22.4 (4.1)† Binging = 2.3 (1.8)‡ Purging = 0.6 (1.3)‡ EDI (DT) = 8.5 (5.0)‡ EDI (bulimia) = 5.2 (4.7)‡	N/A	N/A	N/A
Total		134	130									
Average		45	43	95	23							
SD		37	38	9	7							

Note: BMI, body mass index; EAT, Eating Attitudes Test; EDI, Eating Disorder Inventory (DT, drive for thinness subscale); EDE-Q, Eating Disorder Examination; EDES, Eating Disorder Evaluation Scale; EBW, expected body weight; IBW, ideal body weight; N/A, not applicable. This table describes the discharge and follow-up (FU) data for the RCTs in this review. The significance levels for outcome data are indicated by the following: *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$. The following indicates significant differences in PHP data when compared with the study's control group: †$p < 0.01$. ‡$p < 0.001$.

varied widely. The shortest reported average treatment duration was 22.0 ± 13.4 days (Treat et al., 2008), whereas the longest reported average treatment duration was 182 days (*SD* not reported) (Exterkate et al., 2009). The remaining PHP studies in the review assessed patients that were treated for an average of at least 10 weeks, with the majority falling between 10 and 16 weeks (Crino & Djokvucic, 2010; Dancyger et al., 2003; Goldstein et al., 2011; Henderson et al., 2014; Herpertz-Dahlmann et al., 2014; Jones et al., 2007; Kong, 2005; Manara et al., 2005; Olmsted et al., 2003; Olmsted et al., 2013; Ornstein et al., 2012; Schaffner & Buchanan, 2008; Zeeck et al., 2004; Zeeck et al., 2011).

Residential treatment programmes

The residential treatment programmes discussed in this review (Brewerton & Costin, 2011a, 2011b Delinsky et al., 2010; Hoffart, Lysebo, Sommerfeldt, & Rø, 2010; Lowe et al., 2003; McHugh, 2007; Weltzin et al., 2014) are 24-hour care facilities designed specifically for patients with ED that do not require continuous medical monitoring, all of which primarily used CBT and/or dialectical behaviour therapy in addition to other treatment modalities (e.g. interpersonal, psychodynamic, art therapy). These studies included programmes with a multidisciplinary treatment approach including individual, family and group therapy, nutrition counselling, nursing and medication management. Patients admitted to residential treatment programmes were monitored 24 hours a day. Meals were highly structured, and patients were required to attend various therapy groups daily. Of the studies reviewed, treatment duration varied. The shortest reported average treatment duration was 28.5 ± 12.3 days (McHugh, 2007), whereas the longest reported average treatment duration was 105 days (*SD* not reported) (Hoffart et al., 2010).

Analysis of outcome data

The studies were divided into two groups: first, open trial studies that reported outcome at the time of discharge from treatment (Table 1) and, second, the open trial studies that reported outcome at some time interval after discharge (Table 2). The three randomized controlled trials (RCTs), two of which includes follow-up, are reported in Table 3 and are discussed in detail.

Outcomes at time of discharge from PHP treatment

Adults

Ten open PHP trials focused primarily on adults (Abbate-Daga et al., 2015; Crino & Djokvucic, 2010; Exterkate et al., 2009; Fittig et al., 2008; Goddard et al., 2013; Jones et al., 2007; Olmsted et al., 2003; Olmsted et al., 2013; Willinge et al., 2010; Zeeck et al., 2004) (Table 1). All of these studies reported significant improvements in BMI for patients with AN and/or reductions in binge and purge frequency for patients with BN; six of these studies reported medium to large effect sizes (Abbate-Daga et al., 2015; Crino & Djokvucic, 2010; Exterkate et al., 2009; Fittig et al., 2008; Willinge et al., 2010; Zeeck et al., 2004). For psychological measures, several studies reported significant improvements in measures of depression and anxiety, in addition to the specific measures of ED pathology (Abbate-Daga et al., 2015; Crino & Djokvucic, 2010; Exterkate et al., 2009; Fittig et al., 2008; Jones et al., 2007; Olmsted et al., 2003; Olmsted et al., 2013; Willinge et al., 2010). Zeeck et al. (2004) provided additional support for PHP treatment for BN by matching a PHP sample to an extended inpatient sample; no significant differences between the two modalities in treating patients were reported, as measured by the Eating Disorder Inventory (EDI).

Mixed sample programmes

One uniquely designed study (Treat et al., 2008), which investigated both adolescents and adults with AN in a transition from inpatient to PHP treatment, reported that 35% of those who completed inpatient treatment for 5 weeks, followed by 3 weeks of PHP treatment, demonstrated an 'excellent' outcome ('>90% ideal BMI at discharge and not losing >0.15 kg/week and no regular use of compensatory measures in last week of program'). Other findings support that age at admission does not influence treatment outcome in PHP settings (Dancyger et al., 2003). In Dancyger et al.'s study including adolescents and adults with AN, BN and EDNOS, both the adolescents' and adults' per cent ideal body weight (%IBW) increased after an average of 15 weeks of multidisciplinary treatment. However, significance levels were not discussed. More recent data (Manara et al., 2005) report significant results, showing that both adult and adolescent patients with AN had an increased BMI after PHP treatment. This study also showed that both the patients with AN and those with BN reported significant improvement across most EDI-2 subscales at discharge.

Adolescent and family programmes

Six studies evaluating PHP treatment for adolescents included families, to some extent, within the treatment protocol. Four of these studies (Girz et al., 2013; Henderson et al., 2014; Hoste, 2015; Ornstein et al., 2012) evaluated FBT. Henderson et al., Ornstein et al. and Hoste all reported significant improvements in BMI for low-weight patients. A significant improvement in ED symptoms was also seen across all three of these studies for all patient groups. The fourth FBT study (Girz et al., 2013) took a different approach to analysis, measuring outcomes at different time points throughout treatment. A significant decrease in ED symptoms in adolescents was achieved after 6 months of treatment. Additionally, the 6 months of treatment resulted in an average %IBW greater than 95%; however, significance levels were not discussed for weight. Two other studies (Goldstein et al., 2011; Schaffner & Buchanan, 2008) examined programmes that included families but to a lesser extent than the traditional FBT. However, they also reported significant improvements in ED behaviours and weight.

Randomized controlled trials

Three RCTs that explored PHP outcome were identified. Two studies compared PHP patients with a control group in extended inpatient treatment (Herpertz-Dahlmann et al., 2014; Zeeck et al., 2011), while the other study compared PHP patients with a control group receiving traditional outpatient therapy (Kong, 2005). As RCTs provide stronger evidence than open trials, each of these studies is discussed separately.

Herpertz-Dahlmann et al. (2014) conducted a multicentre trial in which adolescents with AN who were hospitalized for 3 weeks

were randomly assigned to continue inpatient treatment (control group) or step down to PHP treatment. One hundred seventy-two patients with AN were randomly assigned to treatment: 85 to inpatient and 87 to PHP. Patients with AN from six centres in Germany, aged 11–18 years, were eligible if they had a BMI below the 10th percentile and it was their first hospital admission for AN. The treatment programme and intensity in both study groups were identical. Patients were discharged when target weight was reached; inpatients averaged 14.6 (±6.0) weeks of treatment, while PHP patients averaged 16.5 (±7.0) weeks of treatment. Outcomes suggest that PHP was equivalent to continued inpatient treatment, with similar levels of weight gain and improvement in ED symptoms in both groups. Furthermore, PHP patients reported greater mental well-being and psychosexual adjustment.

Zeeck et al. (2011) examined 43 adult patients with BN, whose symptoms did not remit in 25 sessions of outpatient therapy over 2 years. Patients were randomly assigned to either PHP or inpatient (control group) treatments. Global ED pathology—as indexed by the EDI—and binge frequency improved significantly in both groups, without between-group differences. Furthermore, 41% of PHP patients showed a complete remission at time of discharge, compared with 33% of inpatients.

Kong (2005) compared a PHP treatment group with a traditional outpatient therapy control group, including adults with AN, BN and EDNOS. Volunteers from an outpatient clinic were randomly assigned to one of the two groups. For both groups, treatment duration was determined for each patient individually and lasted from 8 to 14 weeks. Results favoured PHP; patients with AN from the PHP group displayed significantly greater improvements in weight gain when compared with those receiving outpatient therapy. Further, PHP patients with BN experienced a greater decrease in binging and purging frequencies when compared with the outpatient control group. Lastly, several measures of psychological symptoms showed greater improvement after PHP in comparison with the outpatient controls.

Outcomes at time of discharge from residential treatment

Adults

Two residential treatment studies included an adult-only sample, and both revealed improvements across most outcome measures (Brewerton & Costin, 2011a, 2011b, Hoffart et al., 2010) (Table 1). One of the studies evaluated patients with BN and reported a significant reduction in binge frequency, without a reduction in purge frequency, although specific scores were not provided, making the finding difficult to interpret (Hoffart et al., 2010).

Mixed sample programmes

Two residential treatment studies reported data collected from mixed adolescent and adult samples (Delinsky et al., 2010; Lowe et al., 2003). When evaluating low-weight patients, both studies found significant increases in BMI or per cent expected body weight (%EBW). They also both found significant improvements in self-report measures of eating pathology when comparing intake scores with discharge scores.

Adolescents

Two residential treatment studies focused primarily on adolescents (McHugh, 2007; Weltzin et al., 2014). With regard to low weight patients, both studies reported overall improvements in BMI after treatment. However, one study did not indicate if the results were significant (Weltzin et al., 2014). In terms of psychological symptoms, both studies reported significant improvements in ED symptoms as measured by the Eating Disorder Examination (EDE-Q) (Weltzin et al., 2014) and the EDI-2 (McHugh, 2007).

Comparison of discharge outcome data for PHP and residential treatment open trials

A total of 19 PHP open trials examined outcome at discharge. A total of 3315 (an average of 174 ± 324 per study) patients entered treatment, and a total of 2664 (an average of 140 ± 268 per study) patients provided data at time of discharge. Thus, 80% of the patients from the initial samples had discharge data. For the 18 studies that reported duration of treatment, the average was 93 ± 46 days. The average age of participants was 21 ± 5 years.

A total of six residential open trials examined outcome at discharge. A total of 1452 (an average of 242 ± 337 per study) patients entered treatment, and a total of 927 (an average of 155 ± 168 per study) patients provided data at time of discharge. Thus, 64% of the patients from the initial samples had discharge data. All six studies reported treatment duration; the average was 58 ± 33 days. The average age of the participants was 23 ± 7 years.

A comparison between the PHP and residential programmes, using two-tailed *t*-tests, revealed no significant differences in initial sample number ($t = 0.44$, $p = 0.66$), discharge number ($t = 0.12$, $p = 0.90$), per cent with discharge data ($t = 0.83$, $p = 0.42$), age ($t = 0.52$, $p = 0.61$) or duration of treatment ($t = 1.76$, $p = 0.09$). For PHP, all but one study (Dancyger et al., 2003) reported a significant improvement in weight and/or ED behaviour at time of discharge. All residential programmes reported an improvement in weight and/or behaviour at time of discharge.

Analysis of follow-up data at some interval after discharge

While specific definitions of recovery vary between studies, a robust consensus among the field exists in that recovery is best gauged over time (Carter, Blackmore, Sutandar-Pinnock, & Woodside, 2004; Field et al., 1997), with multiple relapses often limiting the extent to which illness status at discharge can reliably predict illness status over time. Thus, a more meaningful indicant of treatment efficacy is afforded when the trajectory of treatment outcome is assessed over time, necessitating the reporting of follow-up data. However, many of the reviewed studies did not include follow-up assessments. Of the 14 eligible studies that provided follow-up data, 11 examined PHP settings. Two of these were RCTs and are discussed in the next paragraphs. Three open trial follow-up studies for residential settings were identified.

In considering the utility of follow-up data, it is important to note the response rate at follow-up, with low response rates

reducing the extent to which treatment efficacy can be measured. Follow-up periods varied greatly amongst studies and often across different patients within the same study. The duration from discharge to follow-up ranged from 3 months (Lowe et al., 2003; Willinge et al., 2010; Jones et al., 2007) to 10 years (Brewerton & Costin, 2011b).

Outcome at follow-up for PHP treatment

Of the nine PHP studies providing follow-up data, there was a wide discrepancy in response rate at follow-up ranging from 43% (Fittig et al., 2008) to 99% (Herpertz-Dahlmann et al., 2014) (Table 2). Furthermore, not all patients in each study completed every follow-up measure.

Adults

Five open trial studies focusing primarily on adult outcomes included follow-up data. Three of them found patient status at discharge to be maintained or improved at follow-up (Abbate-Daga et al., 2015; Fittig et al., 2008; Willinge et al., 2010). Both Willinge et al. (2010) and Abbate-Daga et al. (2015) reported that weight gain in patients with AN measured at discharge was maintained at follow-up, while psychological symptoms improved further. Moreover, Fittig et al. (2008) found that the improvements in weight for patients with AN, the reduction in binging and purging in patients with BN and the decreases in drive for thinness in both groups that were reached at discharge were all further improved at 18-month follow-up. Effect size was exceptionally large (2.65) for BMI in the AN-restricting group. Jones et al. (2007) reported only trends because of the difficulty in collecting follow-up data. Of the 15 patients that were included in follow-up analysis, 12 showed continued improvement in ED symptomatology, while 8 showed improvements in mood (Jones et al., 2007). Similarly, Zeeck et al. (2004) lacked significant results but also found a trend at the one-and-a-half year follow-up, suggesting further improvement in ED symptoms in patients with BN after discharge.

Mixed sample programmes

Treat et al. (2008) examined the results of inpatient treatment followed by PHP treatment, as discussed previously. The trial's 6-month follow-up status was deemed 'successful' if the patient maintained outpatient status in the specific treatment system. Approximately half of the patients with a 'good' outcome ('85–90% ideal BMI, not losing >0.15 kg/week, and no compensatory measures in last week of program') at discharge were able to maintain their outpatient treatment status for 6 months rather than regressing to require a higher level of care. The other half of the patients was referred to a higher level of care. Manara et al. (2005) reported assessment at 6-month follow-up, revealing that improved BMI was maintained in patients with AN. Moreover, the reductions in psychological symptoms for both the patients with AN and those with BN were maintained at follow-up assessment.

Adolescent and family programmes

Two adolescent PHP studies reported follow-up data collected at 6 months after discharge (Goldstein et al., 2011; Henderson et al., 2014). Examining weight gain in low-weight patients, Henderson et al. (2014) found that 65% of patients had a healthy BMI (defined in the study as BMI > 19%) at follow-up. While this was a decrease from the 87% with a healthy BMI at discharge, the average BMI at follow-up was still significantly greater than that before treatment. Goldstein et al. (2011) found the improvements in weight gain measured at discharge to be maintained at the 6-month follow-up. When comparing pretreatment BMI with that at 6-month follow-up, this change reflects a large effect size of 0.86. Both studies also found further improvements in psychological measures at follow-up, compared with scores at discharge.

Outcome at follow-up for residential treatment

Three residential studies that met inclusion criteria reported follow-up data (Brewerton & Costin, 2011b; Lowe et al., 2003; Weltzin et al., 2014) (Table 2). Although none of these studies were RCTs, all reported that improvements in BMI and eating pathology were maintained at follow-up. However, a large discrepancy in follow-up period between these studies precludes a meaningful comparison across studies. Thus, in considering the barriers to a cross-study comparison, each of these studies is discussed individually.

Brewerton and Costin (2011a) reported that at follow-up assessment, which for AN averaged 4.6 ± 3.1 years after discharge, 70% of the low-weight patients demonstrated weight recovery, defined as a BMI > 18. Patients with BN completed their follow-up assessments at an average of 3.8 ± 2.7 years after discharge, at which point 62% of the sample had achieved a 'good' outcome (Morgan & Russell, 1975). Improvements from admission to follow-up were also found in almost all of the EDI-2 subscales for both patient groups. It is important to note, for this study in particular, the significant variability in the follow-up intervals within the participants, ranging from 1 to 10 years.

Lowe et al. (2003) conducted a 3-month follow-up after residential treatment and found that improvements measured at discharge were maintained at follow-up. BMI in low-weight patients was significantly improved when compared with pretreatment data. Significant psychological improvements for patients with AN and BN were also noted when compared with pretreatment scores. Interestingly, a significant inverse relationship between the degrees of change during and after treatment was found; the more patients improved in treatment, the more they regressed or stopped improving outside of treatment.

Lastly, Weltzin et al. (2014) reported that for AN, the increase in BMI from admission to discharge was maintained at follow-up. However, because of the limited number of survey respondents, the significance of these data was unclear. A sustained improvement in EDE-Q scores at 24-month follow-up, however, was found to be significant.

Comparison of outcome at follow-up for PHP and residential treatment open trials

A total of nine PHP open trials included follow-up data. A total of 659 (an average of 73 ± 81 per study) patients entered treatment, and a total of 380 (an average of 42 ± 33 per study) patients were followed up at some time point after discharge. Thus, there was an average per cent follow-up (% FU) per study of 66 ± 20% (see % FU in Table 2). For the nine studies, the average time to follow-up was 9 ± 6 months.

A total of three residential open trials included follow-up data. A total of 1280 (an average of 427 ± 426 per study) patients entered treatment, and a total of 319 (an average of 106 ± 45 per study) patients were followed up at some time point after discharge. Thus, there was an average % FU per study of 37 ± 20% (see % FU in Table 2). The average time to follow-up was 26 ± 24 months.

A comparison between the PHP and residential programmes that reported follow-up data, using two-tailed *t*-tests, revealed a significantly larger initial sample size for the residential studies ($t = 2.60$, $p = 0.03$). The number of participants followed up per study for residential treatment was also significantly greater (2.68, $p = 0.02$). However, the PHP studies had a higher average % FU, per study, than residential studies ($t = 2.19$, $p = 0.05$). Lastly, there was a trend showing that residential studies waited a significantly longer time ($t = 2.17$, $p = 0.06$) before follow-up data were collected.

For PHP, six out of the nine studies reported significant improvement in weight and/or ED behaviour at time of follow-up when compared with pretreatment data. All residential programmes reported an improvement in weight and/or behaviour at time of follow-up.

Outcome at follow-up for randomized controlled trials

Two RCTs followed up after discharge, both of which were PHP studies (Table 3). Because RCTs provide strong evidence, these two studies are discussed separately.

Zeeck et al. (2011) conducted a 3-year follow-up assessment for patients with BN. The authors expanded on the findings from their 2009 paper (Zeeck et al., 2009). Two patient groups were evaluated: one that completed PHP and one that completed inpatient treatment. At 3 years post discharge, 72% of PHP patients were in partial or complete remission and 69% of the inpatient group was in either partial or complete remission, demonstrating that PHP was not inferior to inpatient treatment. Moreover, there was a statistically significant improvement in EDI scores at follow-up.

Herpertz-Dahlmann et al. (2014) compared PHP after short inpatient care with continued inpatient care for adolescent females, as discussed earlier. For follow-up analysis, the primary outcome was the increase in BMI from admission to 12-month follow-up, adjusted for age and duration of illness (non-inferiority margin of 0.75 kg/m^2). PHP was similar to inpatient treatment with respect to BMI at 12-month follow-up. Follow-up was attained for 161 patients (94% of the sample). The authors concluded that PHP after short inpatient care in adolescent patients with nonchronic AN was similar to inpatient care for weight restoration and maintenance during the first year after admission. Thus, PHP might be a safe and less costly alternative to inpatient treatment.

Discussion

Overview

The purpose of this paper is to review ED outcome studies for PHP and residential programmes in the interval since the Zipfel et al. (2001) review. There were three categories of outcome studies. The largest number of studies (Table 1) report outcome at the time of discharge. These studies tended to be open trials and almost all reported improvement in weight and/or behavioural symptoms at the time of discharge. However, these findings are not surprising, because discharge is typically not recommended until the patient is within 5–10% of %EBW and demonstrates improved functional behaviour with a significant decrease in symptoms (Halmi, 2007). Given the serious and life-threatening nature of ED, it can be argued that reversing severe emaciation and the nutritional deficits of AN by the time of discharge has a critical value. This may result in reduced morbidity and mortality, even if relapse occurs.

While the findings reported at discharge appear favourable for both PHP and residential treatment settings, the lack of long-term follow-up data limits the clinical utility of these findings. For instance, recent evidence shows that patient status at the end of treatment does not necessarily predict recovery status at longer-term follow-up (Lock et al., 2013), suggesting that relapse is possible for many patients deemed recovered at the end of treatment. Similarly, recent data pertaining to adolescents engaged in FBT suggest that while 50–70% may be weight restored by the end of treatment (Le Grange & Eisler, 2009), only a third remain weight restored at 4-year follow-up (Le Grange et al., 2014), further suggesting that treatment outcome at discharge may provide a skewed picture of longer-term symptom remission. Thus, indexing the efficacy of treatment settings based primarily on outcome data provided at discharge may lend itself to spurious findings; this may be a particular risk in intensive treatment settings when closely monitored staff-driven symptom remission is a common prerequisite for discharge. For instance, one residential treatment study (Lowe et al., 2003) suggested that greater patient improvement resulted in a greater degree of symptom relapse beyond discharge. This illustrated a unique association between programme-based progress and outpatient-based symptom relapse.

There were 12 uncontrolled trials (Table 2) that offered follow-up data at some interval after discharge. Six of the nine PHP studies in this group and all three residential studies reported significant improvement in weight and/or ED behaviour at discharge when compared with baseline data. However, it is important to note that the proportion of patients followed by PHP and residential programmes was problematic. Still, the average percentage of follow-up data for the open trials for PHP was significantly greater (66%) than that for residential treatment (37%). The loss of follow-up data raises questions about conclusions, particularly for the residential studies. Few studies utilized missing data management strategies. The exclusion of treatment noncompleters from analyses may significantly obscure indices of treatment efficacy. Empirical comparisons of treatment completers versus noncompleters reveals greater symptomatology amongst those who drop out, in both AN and BN samples (Fittig et al., 2008). Thus, an overrepresentation of favourable treatment outcomes is more likely when excluding noncompleters from analyses (Pike, 1998; Vandereycken & Pierloot, 1983). However, it is important to emphasize that many studies we reviewed, particularly PHP studies reporting outcome at discharge, did report very low drop-out rates (Crino & Djokvucic, 2010; Girz et al., 2013; Goddard et al., 2013; Goldstein et al., 2011; Henderson et al., 2014; Ornstein et al., 2012; Treat et al., 2008; Zeeck et al., 2004).

The most rigorous data were provided by the two RCTs that reported long-term follow-up. But the overall number of patients evaluated through RCTs was small (N = 134). And these studies (Herpertz-Dahlmann et al., 2014; Zeeck et al., 2011) only compared PHP groups with inpatient control groups. However, these studies did report very high rates of follow-up. Herpertz-Dahlmann et al. (2014) followed up with 99% of PHP patients 1 year after discharge, while Zeeck et al. (2011) had complete data sets for 74% (but had some follow-up data for 95%) of PHP patients at 3 years after discharge. Notably, these studies suggest that PHP is similar to inpatient treatment with regard to maintaining symptom remission.

Limitations and directions for further research

In terms of the EPHPP quality assessment of each study, none of the studies we reviewed achieved a 'strong' global rating. Even with three RCTs, data collectors were not able to successfully implement blinding, preventing any of the studies from receiving a score greater than 'moderate'. Thus, this review highlights the lack of rigorously designed, controlled studies characterizing long-term outcome for these treatment approaches. Although there were three RCTs for PHP treatment supporting its efficacy, there were no controlled studies for residential treatment. It was therefore not possible to conduct a meta-analysis on this body of data. Instead, we chose a narrative type of approach that describes the current literature, with the hope that this will stimulate the field to advance our understanding of the efficacy of these approaches.

Some of the other problems we observed include inconsistency between studies. Sample size—which ranged from 16 (Goddard et al., 2013) to 917 (Lowe et al., 2003)—varied. Additionally, outcome measures differed between studies. In certain cases, some assumptions had to be made in order to provide comparisons. For example many different types of assessment instruments were included, and averages sometimes had to be computed from the data provided. Moreover, studies at follow-up tended to use self-report data, rather than observer data, or, in some cases, did not indicate whether BMI was observed or self-reported. Treatment duration also varied significantly in different settings, which is likely because of insurance and financial constraints, but of course limits our ability to compare studies with each other. Another limitation is that both residential and PHPs vary widely in content and treatment models, and this limits the generalizability of findings. The lack of detailed information regarding treatment approach was among the reasons that it was not possible to determine whether treatment approach had an effect on outcome.

Another issue is the inherent difficulty in obtaining empirically controlled data, with the random allocation of patients to clinical conditions being superseded by the clinical need to match patients to the level of care most appropriate for the severity of their illnesses. Strictly controlled data for the treatment of ED across intensive levels of care is, therefore, understandably sparse. Moreover, there is a need to comprehensively explicate the mediators and moderators of treatment efficacy across a transdiagnostic spectrum of ED in these treatment settings. These weaknesses further highlight the limitations in the current evidence base for ED outcome studies. We recommend additional RCTs for both residential and PHPs and strongly suggest that these studies include follow-up data at time points after discharge.

Conclusion

While acknowledging the noteworthy preliminary evidence, much remains to be learned about the long-term efficacy of PHP treatment and, particularly, residential treatment. Patients in both PHP and residential treatment showed improvement at time of discharge, which is to be expected. Residential treatment results are very preliminary, limited in number of studies available, and difficult to interpret because of dropout rates and limitations of the study samples to Anglo-Saxon countries. What is critically important is determining whether patients relapse after discharge from treatment. Assessment done at the time of discharge may not predict long-term outcome. Most of these studies did not assess outcome at some time point after discharge. Those follow-up studies that have been conducted suffer from a substantial dropout rate, particularly for residential treatment. Thus, these studies may not capture people who have poor outcome and, as a result, may inflate data showing positive response to treatment. Additionally, RCTs have been conducted only for PHP treatment. The two studies that included follow-up had good follow-up rates and showed that PHP is similar to inpatient treatment in terms of efficacy. At the least, this makes the argument that PHP, which is likely a less expensive intervention than inpatient treatment, is equally as successful in treating ED.

While it is likely that residential and PHP treatments provide a very valuable service, there is little in the way of standardized guidelines for the clinical management for those with ED in these programmes. Thus, patients and their families may be exposed to many unproven and potentially non-useful approaches, rather than have data they can rely upon to guide them towards effective therapies. Furthermore, these treatment settings are costly and are becoming increasingly commonplace in the USA, although little evidence to date has documented the cost-effectiveness of residential as compared with PHP treatment, or either approach as compared with other treatment platforms.

The ED field should advance treatment by understanding what works and what does not work in terms of restoring health during the treatment process and reducing relapse after discharge. This requires carefully designed, controlled, large-scale trials. However, finding the means to pay for such studies is challenging in the current funding climate. In the USA, the ED field has recently seen a change in that large corporations now own many of the higher level of care programmes. However, few of these companies have made any investment in scientifically based studies that seek to develop more effective therapies or demonstrate that their programmes improve outcome. If the field began to devote a small percentage of gross revenue to research and development of treatment efficacy, this would help support the type of rigorous studies necessary to improve treatment and outcome.

REFERENCES

Abbate-Daga, G., Gramaglia, C., Preda, S., Comba, E., Brustolin, A., & Fassino, S. (2009). Day hospital programmes for eating disorders: A review of the similarities, differences and goals. *Eating and Weight Disorders*, *14*, e31–e41. 19934635.

Abbate-Daga, G., Marzola, E., De-Bacco, C., Buzzichelli, S., Brustolin, A., Campisi, S., et al. (2015). Day hospital treatment for anorexia nervosa: A 12-month follow-up study. *European Eating Disorders Review*, *23*, 39–398. 10.1002/erv.2369.

Berkman, N., Lohr, K., & Bulik, C. (2007). Outcomes of eating disorders: A systematic review of the literature. *International Journal of Eating Disorders*, *40*, 292–309. 17370291.

Brewerton, T., & Costin, C. (2011a). Long-term outcome of residential treatment for anorexia nervosa and bulimia nervosa. *Eating Disorders*, *19*, 132–144. 10.1080/10640266.2011.551632.

Brewerton, T., & Costin, C. (2011b). Treatment results of anorexia nervosa and bulimia nervosa in a residential treatment program. *Eating Disorders*, *19*, 117–131. 10.1080/10640266.2011.551629.

Carter, J., Blackmore, E., Sutandar-Pinnock, K., & Woodside, D. (2004). Relapse in anorexia nervosa: A survival analysis. *Psychological Medicine*, *34*, 671–679. 15099421.

Crino, N., & Djokvucic, I. (2010). Cohesion to the group and its association with attendance and early treatment response in an adult day-hospital program for eating disorders: A preliminary clinical investigation. *Clinical Psychologist*, *14*, 54–61. 10.1080/13284207.2010.500308.

Dancyger, I., Fornari, V., Schneider, M., Fisher, M., Frank, S., Goodman, B., et al. (2003). Adolescents and eating disorders: An examination of a day treatment program. *Eating and Weight Disorders*, *8*, 242–248. 14649790.

Dejong, H., Broadbent, H., & Schmidt, U. (2012). A systematic review of dropout from treatment in outpatients with anorexia nervosa. *International Journal of Eating Disorders*, *45*(5), 635–647. 10.1002/eat.20956.

Delinsky, S., St. Germain, S., Thomas, J., Craigen, K., Fagley, W., Weigel, T., et al. (2010). Naturalistic study of course, effectiveness, and predictors of outcome among female adolescents in residential treatment for eating disorders. *Eating and Weight Disorders*, *15*, 127–135. 21150248.

Exterkate, C., Vriesendorp, P., & de Jong, C. (2009). Body attitudes in patients with eating disorders at presentation and completion of intensive outpatient day treatment. *Eating Behaviors*, *10*, 16–21. 10.1016/j.eatbeh.2008.10.002.

Field, A., Herzog, D., Keller, M., West, J., Nussbaum, K., & Colditz, G. (1997). Distinguishing recovery from remission in a cohort of bulimic women: How should asymptomatic periods be described? *Journal of Clinical Epidemiology*, *50*, 1339–1345. 9949937.

Fittig, E., Jacobi, C., Backmund, H., Gerlinghoff, M., & Wittchen, H. (2008). Effectiveness of day hospital treatment for anorexia nervosa and bulimia nervosa. *European Eating Disorders Review*, *16*, 341–351. 18668700.

Frisch, M., Franko, D., & Herzog, D. (2006). Arts-based therapies in the treatment of eating disorders. *Eating Disorders*, *14*, 131–142. 16777810.

Girz, L., LaFrance Robinson, A., Foroughe, M., Jasper, K., & Boachie, A. (2013). Adapting family-based therapy to a day hospital programme for adolescents with eating disorders: Preliminary outcomes and trajectories of change. *Journal of Family Therapy*, *35*, 102–120. 10.1111/j.1467.2012.00618x.

Goddard, E., Hibbs, R., Raenker, S., Salerno, L., Arcelus, J., Boughton, N., et al. (2013). A multi-centre cohort study of short term outcomes of hospital treatment for anorexia nervosa in the UK. *BMC Psychiatry*, *13*, 287. 10.1186/1471-244X-13-287.

Goldstein, M., Peters, L., Baillie, A., McVeagh, P., Minshall, G., & Fitzjames, D. (2011). The effectiveness of a day program for the treatment of adolescent anorexia nervosa. *International Journal of Eating Disorders*, *44*, 29–38. 20063371.

Grilo, C., Pagano, M., Stout, R., Markowitz, J., Ansell, E., Pinto, A., et al. (2012). Stressful life events predict eating disorder relapse following remission: Six-year prospective outcomes. *International Journal of Eating Disorders*, *45*, 185–192. 21448971.

Halmi, K. (2007). Management of anorexia nervosa in inpatient and partial hospitalization settings. In Clinical manual of eating disorders (pp. 113–125). Arlington, VA: American Psychiatric Publishing, Inc..

Henderson, K., Buchholz, A., Obeid, N., Mossiere, A., Maras, D., Norris, M., et al. (2014). A family-based eating disorder day treatment program for youth: Examining the clinical and statistical significance of short-term outcomes. *Eating Disorders*, *22*, 1–18. 10.1080/10640266.2014.857512.

Hepburn, Z., & Wilson, K. (2014). Effectiveness of adult day treatment for eating disorders. *Mental Health Review Journal*, *19*, 131–144. 10.1108/MHRJ-01-2013-0003.

Herpertz-Dahlmann, B., Schwarte, R., Krei, M., Egberts, K., Warnke, A., Wewetzer, C., et al. (2014). Day-patient treatment after short inpatient care versus continued inpatient treatment in adolescents with anorexia nervosa (ANDI): A multicentre, randomised, open-label, non-inferiority trial. *Lancet*, *383*, 1222–1229. 10.1016/S0140-6736(13)62411-3.

Hoffart, A., Lysebo, H., Sommerfeldt, B., & Rø, O. (2010). Change process in residential cognitive therapy for bulimia nervosa. *European Eating Disorders Review*, *18*, 367–375. 10.1002/erv.980.

Hoste, R. (2015). Incorporating family-based therapy principles into a partial hospitalization programme for adolescents with anorexia nervosa: Challenges and considerations. *Journal of Family Therapy*, *37*, 41–60. 10.1111/1467-6427.12055.

Jones, A., Bamford, B., Ford, H., & Schreiber-Kounine, C. (2007). How important are motivation and initial Body Mass Index for outcome in day therapy services for eating disorders? *European Eating Disorders Review*, *15*, 283–289. 17676698.

Kaplan, A., Olmsted, M., Carter, J., & Woodside, B. (2001). Matching patient variables to treatment intensity: The continuum of care. *The Psychiatric Clinics of North America*, *24*, 281–292. 11416928.

Kaye, W., Enright, A., & Lesser, S. (1998). Characteristics of eating disorders programs and common problems with third-party payers. *International Journal of Eating Disorders*, *7*, 573–579. 10.1002/1098-108X(198807)7:4<573::AID-EAT2260070417>3.0.CO;2-G.

Kong, S. (2005). Day treatment programme for patients with eating disorders: Randomized controlled trial. *Journal of Advanced Nursing*, *52*, 5–14. 15941455.

Le Grange, D., & Eisler, I. (2009). Family interventions in adolescent anorexia nervosa. *Child and Adolescent Psychiatric Clinics of North America*, *18*, 159–173. 10.1016/j.chc.2008.07.004.

Le Grange, D., Lock, J., Accurso, E., Agras, W., Darcy, A., Forsberg, S., et al. (2014). Relapse from remission at two- to four-year follow-up in two treatments for adolescent anorexia nervosa. *Journal of the American Academy of Child and Adolescent Psychiatry*, *53*, 1162–1167. 10.1016/j.jaac.2014.07.014.

Lock, J., Agras, W., Le Grange, D., Couturier, J., Safer, D., & Bryson, S. (2013). Do end of treatment assessments predict outcome at follow-up in eating disorders? *International Journal of Eating Disorders*, *46*, 771–778. 10.1002/eat.22175.

Lowe, M., Davis, W., Annuziato, R., & Lucks, D. (2003). Inpatient treatment for eating disorders: Outcome at discharge and 3-month follow-up. *Eating Behaviors*, *4*, 385–397. 15000964.

Maguire, S., Le Grange, D., Surgenor, L., Marks, P., Lacey, H., & Touyz, S. (2008). Staging anorexia nervosa: Conceptualizing illness severity. *Early Intervention in Psychiatry*, *2*, 3–10. 21352125.

Manara, F., Manara, A., & Todisco, P. (2005). Correlation between psychometric and biological parameters in anorexic and bulimic patients during and after an intensive day hospital treatment. *Eating and Weight Disorders*, *10*, 236–244. 16755167.

McHugh, M. (2007). Readiness for change and short-term outcomes of female adolescents in residential treatment for anorexia nervosa. *International Journal of Eating Disorders*, *40*, 602–612. 17610253.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D., & The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Journal of Clinical Epidemiology*, *62*, 1006–1012. 10.1016/j.jclinepi.2009.06.005.

Morgan, H., & Russell, G. (1975). Value of family background and clinical features as predictors of long-term outcome in anorexia nervosa. *Psychological Medicine*, *5*, 355–371. 1197538.

Murray, S., Anderson, L., Rockwell, R., Griffiths, S., Le Grange, D., & Kaye, W. (2015). Adapting family-based treatment for adolescent anorexia nervosa across higher levels of patient care. *Eating Disorders*, *23*, 302–314. 26011736.

Murray, S., Loeb, K., & Le Grange, D. (2014). Indexing psychopathology throughout family-based treatment for adolescent anorexia nervosa: Are we on track? *Advances in Eating Disorders: Theory, Research and Practice*, *2*, 93–96. 10.1080/21662630.2013.831522.

National Collaborating Centre for Methods and Tools. 2008. Quality assessment tool for quantitative studies method. Hamilton, ON: McMaster University. Available: http://www.nccmt.ca/registry/view/eng/15.html.

Olmsted, M., Kaplan, A., & Rockert, W. (2003). Relative efficacy of a 4-day versus a 5-day day hospital program. *International Journal of Eating Disorders*, *34*, 411–449. 14566931.

Olmsted, M., McFarlane, T., Trottier, K., & Rockert, W. (2013). Efficacy and intensity of day hospital treatment for eating disorders. *Psychotherapy Research*, *23*, 277–286. 10.1080/10503307.2012.721937.

Ornstein, R., Lane-Loney, S., & Hollenbeak, C. (2012). Clinical outcomes of a novel, family-centered partial hospitalization program for young patients with eating disorders. *Eating and Weight Disorders*, *17*, e170–e177. 23086252.

Pike, K. (1998). Long-term course of anorexia nervosa: Response, relapse, remission, and recovery. *Clinical Psychology Review*, *18*, 447–475. 9638357.

Schaffner, A., & Buchanan, L. (2008). Integrating evidence-based treatments with individual needs in an outpatient facility for eating disorders. *Eating Disorders*, *16*, 378–392. 10.1080/10640260802370549.

Steinhausen, H. (2009). Outcome of eating disorders. *Child and Adolescent Psychiatric Clinics of North America*, *18*, 225–242. 19014869.

Treat, T., McCabe, E., Gaskill, J., & Marcus, M. (2008). Treatment of anorexia nervosa in a specialty care continuum. *International Journal of Eating Disorders*, *41*, 564–572. 18570196.

Twohig, M., Bluett, E., Torgesen, J., Lensegrav-Benson, T., & Quakenbush-Roberts, B. (2014). Who seeks residential treatment? A report of patient characteristics, pathology, and

functioning in females at a residential treatment facility. *Eating Disorders*, *23*, 1–14. 25298220.

Vandereycken, W., & Pierloot, R. (1983). The significance of subclassification in anorexia nervosa: A comparative study of clinical features in 141 patients. *Psychological Medicine*, *13*, 543–549. 6622608.

Weltzin, T., Kay, B., Cornella-Carlson, T., Timmel, P., Klosterman, E., Kinnear, K., et al. (2014). Long-term effects of a multidisciplinary residential treatment model on improvements of symptoms and weight in adolescents with eating disorders. *Journal Groups Addict & Recover*, *9*, 71–85. 10.1080/1556035X.2014.868765.

Willinge, A., Thornton, C., Olmsted, M., & Touyz, S. (2012). The treatment setting for eating disorders: Day patient treatment. In Eating and its disorders (pp. 360–393). Chichester: Wiley-Blackwell.

Willinge, A., Touyz, S., & Thornton, C. (2010). An evaluation of the effectiveness and short-term stability of an innovative Australian day patient programme for eating disorders. *European Eating Disorders Review*, *18*, 220–233. 10.1002/erv.997.

Wilson, G., Vitousek, K., & Loeb, K. (2000). Stepped care treatment for eating disorders. *Journal of Consulting and Clinical Psychology*, *68*, 564–572. 10965631.

Wiseman, C., Sunday, S., Klapper, F., Harris, W., & Halmi, K. (2001). Changing patterns of hospitalization in eating disorder patients. *International Journal of Eating Disorders*, *30*, 60–74. 11439410.

Zeeck, A., Herzog, T., & Hartmann, A. (2004). Day clinic or inpatient care for severe bulimia nervosa? *European Eating Disorders Review*, *12*, 79–86. 10.1002/erv.535.

Zeeck, A., Weber, S., Sandholz, A., Joos, A., & Hartmann, A. (2011). Stability of long-term outcome in bulimia nervosa: A 3-year follow-up. *Journal of Clinical Psychology*, *67*, 318–327. 21254059.

Zeeck, A., Weber, S., Sandholz, A., Wetzler-Burmeister, E., Wirsching, M., & Hartmann, A. (2009). Inpatient versus day clinic treatment for bulimia nervosa: A randomized trial. *Psychotherapy and Psychosomatics*, *78*, 152–160. 19270470.

Zipfel, S., Reas, D., Thornton, C., Olmsted, M., Williamson, D., Gerlinghoff, M., et al. (2001). Day hospitalization programs for eating disorders: A systematic review of the literature. *International Journal of Eating Disorders*, *31*, 105–117. 11920973.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site.