Preventive Medicine 52 (2011) 152–158

Contents lists available at ScienceDirect

Preventive Medicine

journal homepage: www.elsevier.com/locate/ypmed

Real-world effectiveness of a German school-based intervention for primary prevention of anorexia nervosa in preadolescent girls

Katharina Wick *, Christina Brix [1], Bianca Bormann [1], Melanie Sowa [1], Bernhard Strauss [1], Uwe Berger [1]

Institute of Psychosocial Medicine and Psychotherapy, University Hospital, Friedrich-Schiller-University Jena, Stoystr. 3, 07740 Jena, Germany

ARTICLE INFO

Available online 13 December 2010

Keywords:
Eating disorders
Anorexia
Primary prevention
School based programs

ABSTRACT

Objective. Anorexia nervosa ('AN') is notoriously difficult to treat, has high mortality rates, and has a prevalence peak in 15-year-old girls. We developed a German school-based intervention program ('PriMa') for the primary prevention of AN in preadolescent girls and assessed the effects in a sample of Thuringian girls.

Method. Intervention involved nine guided lessons with special posters and group discussions. A parallel controlled trial with pre-post measurements and a three-month follow-up was conducted in 92 Thuringian schools ($n = 1553$ girls) in 2007 and 2008. Primary outcomes were conspicuous eating behavior, body self esteem, and AN-related knowledge.

Results. After adjusting for the girls' ages and the type of school, we observed significant improvements in the areas of knowledge ($d = .24$) and body self esteem ($d = .29$), but not for eating behavior.

Conclusion. The PriMa intervention provides an efficient and practical model to increase AN-related protection factors.

© 2010 Elsevier Inc. All rights reserved.

Introduction

Anorexia nervosa ('AN') is one of the most serious psychiatric disorders in girls and young women. The Diagnostic and Statistical Manual of Mental Disorders (DSM IV-TR) includes four criteria for diagnosing AN (see Table 1). Because it is difficult to treat (Steinhausen, 2002) and because of its high mortality rate (>15%; Zipfel et al., 2000), the disease has gained much attention from the public and from the scientific community. AN's prevalence is estimated at 0.5 to 3.7%, with a peak of incidence at age 15 (Hoek and van Hoeken, 2003; Bulik et al., 2006). Several of AN's symptoms are considerably more common (Grilo, 2006), which are known as sub-clinical expressions of AN. In view of this, groups around the world have spent decades developing prevention programs to prevent the onset and consequences of AN and other eating disorders (for review, see Levine and Smolak, 2006).

According to the standards of the Society of Prevention Research ('SPR'), a successful prevention approach includes the demonstration of efficacy (level 1) as well as the probation in the field under real-world conditions (level 2 'effectiveness'). 'Real-world conditions' mean regular conditions in everyday life without the motivational character of a new program introduced for the first time. Furthermore, an appropriate dissemination is necessary to secure sustainability (level 3 'broad dissemination'; Flay et al., 2005).

Most of the existing prevention programs do not meet all of the standards, including level 2 and 3. For this reason, in 2003, our institution began cooperating with the Thuringian Ministry of Education ('TME') to develop the program called 'Primary prevention of anorexia nervosa in preadolescent girls' ('PriMa'). Our program focuses exclusively on AN, because it occurs earlier (peak of incidence: 15 years) than bulimia nervosa and binge-eating disorder (peak of incidence: 18 years; Grilo, 2006). In 2004 and 2005, we conducted a pilot study to evaluate the process and impact of PriMa (level-1 evaluation; $n = 1006$ girls; mean age = 12 years; Berger, 2008). This study showed significant improvements in pre–post measurements and pre–follow-up measurement on body self esteem, eating behavior, and knowledge.

Furthermore, we paired with the Heidelberg Prevention Center (www.h-p-z.de; level-3-evaluation; see Berger et al., 2008) to create appropriate distribution structures. Two well-known meta-analyses (Stice and Shaw, 2004; Stice et al., 2007) have summarized the effects of over 80 international programs for the prevention of eating disorders. The authors clearly identify factors that successful programs exhibit:

- interactive (vs. didactic),
- selective, for risk groups (vs. universal),
- multisession (vs. single-session),
- gender specific, offered only to females (vs. co-educative),
- offered to participants over age 15 (vs. younger participants), and
- delivered by professionals (vs. endogenous providers like teachers).

* Corresponding author. Fax: +49 3641 936546.
E-mail addresses: katharina.wick@med.uni-jena.de (K. Wick), christina.brix@med.uni-jena.de (C. Brix), bianca.bormann@med.uni-jena.de (B. Bormann), melanie.sowa@gmx.de (M. Sowa), bernhard.strauss@med.uni-jena.de (B. Strauss), uwe.berger@uni-jena.de (U. Berger).
[1] Fax: +49 3641 936546.

0091-7435/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ypmed.2010.11.022

Table 1
DSM-IV criteria for Anorexia nervosa (AN), and examples for criteria A–C from the PriMa study (Thuringia, Germany, 2007-2008).

A	A refusal to maintain body weight at or above a minimally normal weight for age and height (usually less than 85% of ideal body weight)	The 3rd percentile (weight group: extremely underweight) according to Kromeyer-Hauschild et al. (2001)
B	Intense fear of gaining weight or becoming fat	I am preoccupied with a desire to be thinner.[a] Engage in dieting behavior (eat little or only selected food).[a] I am terrified about being overweight.[a]
C	Disturbance in the way one's body weight or shape is experienced, with denial of current low body weight	I am preoccupied with the thought of having fat on my body.[a]
D	Amenorrhea in postmenarchal females of at least 3 menstrual cycles[b]	

[a] Items included in the AN-specific subscale.
[b] Not relevant for the population.

Because of the early peak of incidence of AN, we decided to focus PriMa on 6th-grade girls with a mean age of 12 years. According to most of the programs, PriMa is geared to the WHO-recommended 'Setting Approach' (Stewart-Brown, 2006), realizing the program in schools. We focused primarily on girls because adolescent girls and young women are 10-times more likely to exhibit incidents of AN (Grilo, 2006). We adopted a universal prevention program to avoid labeling the participating girls who were already suffering from (subclinical) eating disorders or showing specific risk factors. Other universal prevention programs were very effective (Becker et al., 2005; Raich et al., 2009) and primary prevention by definition includes this universal character. Furthermore, our cooperation partners insisted using teachers to implement the project rather than independent professionals on the theory that teachers would improve the program's sustainability.

Existing prevention programs concerning eating disorders focused on decreasing risk factors (e.g. weight concerns, negative body image, dieting, and low social support; Jacobi et al., 2004) and strengthening protective factors (e.g. body self esteem, and knowledge). The present study aims to follow this primary preventive approach (Caplan, 1964) to reduce the occurrence of AN (see Table 2). Furthermore we intend to replicate the results of the pilot study concerning the effects of the PriMa program under real-world conditions (level 2; Berger et al., 2008). In this context, the focused research question is whether the primary preventive intervention program, PriMa, is able to significantly improve the primary outcomes eating behavior, body self-esteem, and AN-related knowledge in comparison to a control group?

Methods

Study design and subjects

We designed a controlled trial with (post-hoc) parallel group assignment using pre-post-design, supplemented by a follow-up-measurement after about three months. Before starting the program, we first obtained the consents of the parents, the supervisor of the TME, and the ethics committee of the Jena University Hospital. Our close cooperation with TME required several compromises because TME was focused on the practical effects of the program rather than the underlying research project. Because the program was open to all interested schools, our cooperation partners insisted on conducting the program without a randomized group assignment. Furthermore, we were prohibited from creating waiting-list conditions and switching classes (intervention group vs. control group) at mid-year. This was necessary to ensure sustainability and to not discourage the participating schools.

Table 2
Overview of the nine PriMa-lections.

PriMa-lection	AN-specific risk factors and *protective factors*	Content of PriMa-lections	Goals of PriMa-lections	Impact on outcomes (*item example*)
1	Thin body ideal *Media literacy*	Beauty ideals Being thin How it is to be a model	Reflecting beauty ideals	Increasing body self esteem and knowledge (*Being thin is always an advantage)*
2	Low level of attention *Knowledge about alternatives to gain attention*	Expectations Rebellion, rebellion Against authorities Compromises	Reflecting thoughts and feelings against parents Discussing attention by emaciation Getting to know possible alternatives	Increasing eating behavior *(I feel that others would prefer if I ate more)* and knowledge
3	Perfectionism *Ability to compensate weaknesses*	Superiority Perfectionism High self-esteem	Reflecting one's own ambitions	Increasing eating behavior *(I am preoccupied with a desire to be thinner*), body self esteem and knowledge
4	Overevaluation of food dieting *Knowledge about healthy nutrition*	Healthy eating effect of waiver	Dealing with healthy and unhealthy foods Training of consumption	Increasing eating behavior *(I avoid foods with sugar in them)* and knowledge
5	Negative body image *Positive body image*	Body image Body awareness	Getting a better body awareness Reflecting attitudes, feelings against one's own body Understanding of distorted body awareness	Increasing body self esteem *(I am satisfied with my own body)* and knowledge
6	Inadequate coping strategies *Adequate coping strategies*	Problems in family Sadness	Management of feelings and moods Reflecting thoughts and feelings against parents	Increasing knowledge *(If there are problems, then it is a waste of time, to look for compromises with the parents)*
7	Inadequate problem-solving *Adequate problem-solving*	Solving problems Recognizing distorted eating behavior	Reflecting own problems Asking for help Learning about alternatives	Increasing knowledge *(If a girl does nothing really well, then she should at least be good in losing weight)*
8	Inadequate coping strategies *Knowledge*	Imprisonment in the addiction Blunting of emotions Balance determines life	Learning how to deal with comparisons Being aware of the danger of dependencies Accepting own deficiencies	Increasing body self esteem *(I am satisfied with my height and weight)* and knowledge
9	Inadequate problem-solving *Knowledge*	Course of anorexia Asking for help Changing roles	Getting to know offers of help/ways out of sadness and depression Being aware of the danger of thinness	Increasing knowledge *(It's never too late to seek help from others)*

In September 2006, we started to invite all Thuringian schools to register 6th-grade teachers for the training session. The assignment to the intervention group ('IG') was carried out by self-selection: 42 schools ($n = 804$) agreed to participate in the project-based lessons, including a written questionnaire. 50 schools agreed to fill out the questionnaire only ($n = 749$) to all three measurement points (control group, 'CG'). Since PriMa was integrated in the regular curriculum, participation was compulsory for all girls.

To rule out systematic differences (= controlling of confounders) between the two groups, the sample was parallelized post hoc after the project implementation, before the analyses (no randomization was possible). First, an age range of 11 to 13 years was chosen as an inclusion criterion, because 10- and 14-year old girls in 6th grade are underrepresented. Subsequently, the schools have been assigned pairwise to adjust differences in school characteristics such as type of school (more regular schools took part), students per class, completeness of the survey period (class level), and measurement time. For these reasons, the sample was reduced to 887 girls. 450 girls participated in the intervention and in accordance 437 served as controls (for details, see Fig. 1).

Intervention

PriMa starts with an information session for parents and a one-day training session for the project-implementing teachers. As recommended by Levine and Smolak (2006), after the training session, the teachers will independently conduct the lessons with the help of a 100-page teaching manual with special workbooks. The training session was conducted by two specially trained and experienced teachers, who our cooperation partners released from regular lessons for the PriMa-project. The 'real' intervention consists of 9 45-minute sessions with optional add-on project-based teaching exercises (e.g. role plays, exercises, and analysis of film sequences). Only the girls of the participating schools took part in the project (the boys were involved in the TOPP program (Teenage Obesity Prevention Program) at the same time, see Schwartze et al., accepted). Each lesson begins with a specially designed poster. On each poster, a Barbie doll in the role of a girl on her way to AN is presented (e.g. first lesson: Barbie is winning a model competition and everyone is acclaiming her). The temporal sequence of the program is shown in Fig. 2, derived from case studies for a better understanding of AN (Gerlinghoff and Backmund, 1993). The authors provide a profound access and unexpected insights concerning patients, their families and the associated social implications. PriMa's nine topics occur consistently in this order in patients suffering from AN. In addition, Table 2 shows the PriMa-specific risk and protective factors and how the specific intervention components were hypothesized to affect outcome.

The methods used in the program focus on dissonance induction (see e.g. Stice et al., 2003; Roehrig et al., 2005) and their resolving by group discussions as well as interactive actions by role plays. These methods deviate from prevention programs with consistently weak effects using psychoeducational content (Stice et al., 2007).

Sample size

To detect a small effect ($d = 0.2$) by Cohen (1988) with a power of 95% and an alpha error of 5% (significance level = $p < .05$; two-tailed) assuming an intraclass coefficient (ICC) of 0.05 (= cluster effect, see Hedges and Hedberg, 2007), we need a sample size of $n = 405$. The power computation was carried out via G-Power 2.0 (Erdfelder et al., 1996) and adjusted by using the inflation formula (Eldridge et al., 2006), assuming an average cluster size of 15 girls per school.

Measuring

The questionnaire contains a total of 67 items. 59 items were used in the present study, including demographic information (school, age, height, weight) and the scales to measure the primary outcomes. Not included in the analyses was the SCOFF questionnaire (5 items), because it was attached for a validation study, only. Furthermore the requested school grades (3 items) were not used in the analyses.

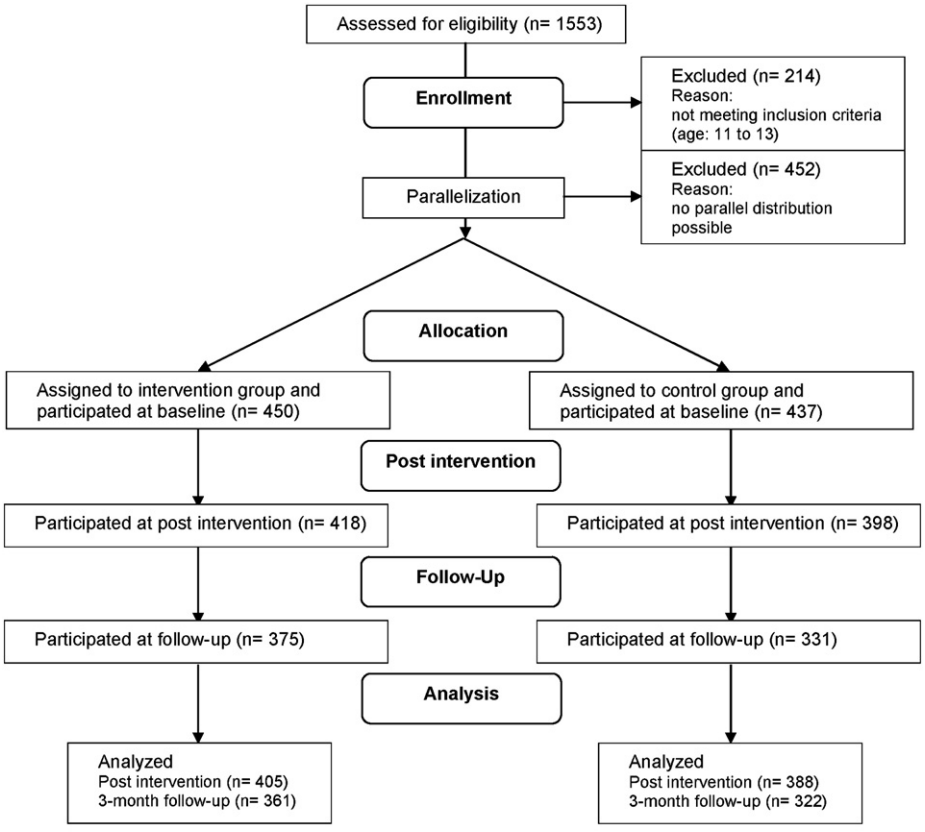

Fig. 1. Study flow diagram.

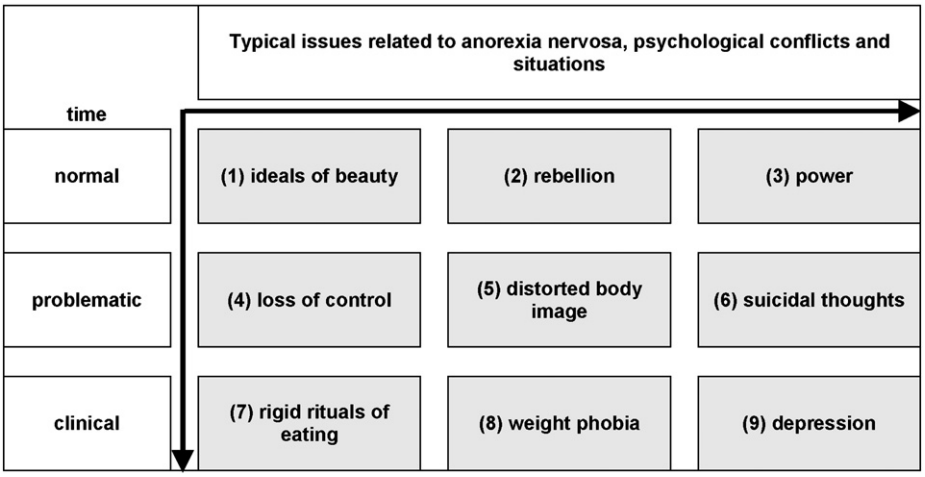

Fig. 2. The "Barbie-Matrix": concept of PriMa including the 9 main topics of the project.

Body self esteem was measured with a scale belonging to the FBeK (Strauss and Richter-Appelt, 1996). It contains 15 items, asking for negative and aggressive reactions towards one's own body and the feeling of attractiveness and identification with one's own body (response categories 'agree' or 'disagree'; scale range 0–15; internal consistency in the present study Cronbach's $\alpha = .89$).

A self-developed scale was used to measure the knowledge about the dangers of AN. For the development of this scale we have chosen from each project lesson one to two questions to repeat what was learned within the AN-specific project (i.e. more females suffer from AN than males). Therefore, the knowledge scale was evaluated by all project collaborators as valid with regards to content and as adequate concerning the requirements. The girls should either affirm the 15 items or deny them or answer with 'don't know' (scale range 0–15; internal consistency in the present study Cronbach's $\alpha = .68$). The item difficulty ranges between 0.40 and 0.90, whereas the item discrimination is on average 0.30.

The girls' eating behavior was assessed with the German-version of the Eating Attitudes Test (Tuschen-Caffier, 2005; original EAT: Garner and Garfinkel, 1979). The questionnaire includes 26 items that represent conspicuous to pathological eating behavior. The answer was given on a six-point scale. The summation of the items was according to the evaluation requirement of the EAT with 0 points for the response categories 'never', 'rarely', and 'sometimes' and one to three points for the response categories 'often', 'very often', and 'always' (scale range 0–78; internal consistency in the present study Cronbach's $\alpha = .89$). Furthermore, the EAT is able to reveal less distinctive disturbances of eating behavior. After Buddeberg-Fischer and Reed (2001) the EAT values can be divided in non- to low-risk (0–9 points) and high-risk ($\geq$10 points) groups. Because PriMa is specifically geared to prevent AN, four items of the EAT-26D were selected and combined in a subscale. 'Eating Disorder Examination' guidelines (EDE; Fairburn and Cooper, 1993) assume that the subscale represents the diagnostic criteria B and C of AN according to the DSM-IV (examples see Table 1). This subscale with a scale range of 0 to 12 achieved in the present sample a good internal consistency of $\alpha = .81$.

To determine the internal validity, the scales were correlated with each other as well as with the girls' self-reported body mass index ('BMI' = weight/height in m 2). These correlations were as follows: eating behavior correlated negatively with body self esteem ($r = -.56$, $p < .01$) and knowledge ($r = -.33$, $p < .01$); BMI correlated positively with 'eating behavior' ($r = .20$, $p < .01$) and negatively with body self esteem ($r = -.46$, $p < .01$). In sum, the reliability and internal validity of the used instruments are satisfactory to good.

Statistical analyses

All calculations were performed with SPSS (version 16.0). For the evaluation we conducted analyses of covariance to control potential baseline differences with 'group' as independent factor, the first measurement point as covariate and the second and third measurement point as dependent variables. To estimate the size of the results, the statistical effect size d (mean difference divided by pooled standard deviation) was calculated. Subgroup analyses for non-risk and risk group followed the same procedure. Supplemental analyses using chi-square tests were carried out to investigate intervention effects on the percentage of extremely underweight participants (equates DSM-IV criterion A, see Table 1).

Results

Patient characteristics

The demographic characteristics of the girls (IG vs. CG) can be seen in Table 3. IG and CG did not differ significantly, except for 'age' ($t(885) = 2.66$, $p = .01$). However, this age difference (12.01 years vs. 11.95 years) has a low substantive importance with an effect of $d = 0.1$ (mean difference about 22 days). Potential distortions through hierarchical effects (individuals in grades in schools) could be ruled out, because the ICC determined from the data is 0.02 (Raudenbush and Bryk, 2002).

793 girls had valid values at the second measurement point concerning the EAT and 683 girls took part at the third measurement

Table 3
Sample characteristics at baseline (Thuringia, Germany, 2007-2008).

Mean scores (SD)	IG ($n = 450$)	CG ($n = 437$)
Age, year	12.0 (0.7)[a]	12.0 (0.5)[a]
Height, cm[b]	158.5 (7.7)	157.8 (7.7)
Weight, kg[b]	48.3 (10.6)	46.9 (10.2)
BMI[b], kg/m^2	19.1 (3.4)	18.7 (3.3)
Age groups, *n* (%)		
11 years	95 (21.1)	68 (15.6)
12 years	235 (52.2)	324 (74.1)
13 years	120 (26.7)	45 (10.3)
Weight groups, *n* (%)		
Underweight (10%: < 10th percentile)[c]	94 (22.3)	102 (25.1)
Normal weight (80%: 10th to 90th percentile)[c]	261 (61.8)	249 (61.3)
Overweight (10%: >90th percentile)[c]	67 (15.9)	55 (13.5)
Type of school, *n* (%)		
Regular school	261 (58)	242 (55.4)
Grammar school	189 (42)	195 (44.6)
EAT group assignment, *n* (%)[d]		
Non-risk group (EAT<10)[e]	289 (64.2)	279 (63.8)
Risk group (EAT$\geq$10)[e]	161 (35.8)	158 (36.2)

[a] Significant difference ($p < .05$).
[b] Self reported; BMI computed by weight and height.
[c] According to Kromeyer-Hauschild et al. (2001).
[d] Classification corresponding to Buddeberg-Fischer and Reed (2001).
[e] EAT = EAT-26D (Eating-Attitudes-Test translated into German).

point. This represents a drop-out rate of 10.6%, respectively 23% (see Fig. 1).

Table 4 lists each mean difference value of all primary outcomes from first to second measurement point and from first to third measurement point and the significant results of the ANCOVAs. The following description refers in detail only to the significant results.

Body self esteem

For body self esteem, a significant group influence (IG vs. CG) could be observed from the first measurement point to the second ($F(1; 764) = 17.10, p < .01, d = 0.29$). Based on the pre-measurement, a higher body self esteem was found at post measurement for the IG (+0.79 points), compared to the CG (+0.05 points).

Knowledge

The changes in knowledge were significantly influenced by the group (IG vs. CG). First, the group affiliation (IG vs. CG) did significantly affect the knowledge at the second measurement point. Girls belonging to the IG showed a greater increase in knowledge (+0.89 points) from first to second measurement point than the knowledge of the girls belonging to the CG (+0.29 points; $F(1; 811) = 15.31$; $p < .01$, $d = 0.22$). These improvements in the IG (+1.12 points) compared to the CG (+0.45 points) could be found for the period from pre- to follow-up measurement, too ($F(1; 701) = 19.61$, $p < .01$, $d = 0.24$).

Subgroup analyses

Differential effects of the intervention for the risk-group (EAT ≥ 10) in comparison to the non-risk-group (EAT < 10) showed significant improvements for body self esteem in the IG compared to the CG from first to second measurement points (risk-group: $F(1; 273) = 7.13$, $p = .01$, $d = .35$; non-risk-group: $F(1; 482) = 8.08$; $p = .01$, $d = .26$). The knowledge was significantly influenced by group assignment (IG vs. CG). This finding occurred from the first to second measurement points (risk-group: $F(1; 282) = 7.92$; $p = ,01$; $d = 0.27$; non-risk-group: $F(1; 510) = 8.04$; $p = ,01$; $d = 0.19$) as well as from first to third measurement point (risk-group: $F(1; 241) = 8.05$; $p = ,01$; $d = 0.28$; non-risk-group: $F(1; 446) = 12.58$; $p < ,01$; $d = 0.22$).

Supplemental analyses

The percentage distribution of extremely underweight participants (<3rd percentile) did not significantly vary between measurement points for IG (pre: 4.2%, post: 4.4%, follow-up: 3.9%) and CG (pre: 4.6%, post: 5.1%, follow-up: 3.1%) nor in comparison of both groups among each other at all three measurement points.

Discussion

The pre–post control group study described above aimed to verify the effectiveness of PriMa against AN for girls attending the 6th grade that are an average of 12 years old. According to Table 4, we found that values for all variables in the IG developed in accordance with the hypotheses. Compared to the CG, a significant intervention effect could be found only for the knowledge at post as well as follow-up measurement. The body self esteem improved significantly only in the short term (pre–post-measurement). For eating behavior, there were no statistically significant changes, neither at post nor at follow-up measurement. After Cohen (1988), the observed significant intervention effects are small. Similar results could be found within the subgroup analyses (risk group vs. non-risk group). A decrease of the percentage of extremely underweight girls was not found, probably because of the short time period of the study. That is why the reduction of the incidence rate of AN and the disease frequency (prevalence) was not investigated in this study, but is part of a larger-scale project in which participants' eating disorder symptoms are assessed in more detail.

The observed improvements within the CG suggest an interpretation of filling out the questionnaire as a mini-intervention. To control this effect in a future study, a so-called baseline-group (Kleimann and Moessle, 2008) should be introduced as a control group with only two measurement points.

Table 4
Mean outcome measures (standard error, standard deviation for strata) for intervention (IG) and control group (CG) at baseline, post intervention and 3-month follow-up (Thuringia, Germany, 2007-2008).

Point of measurement	Baseline		Post intervention[2]		Follow-up[3]	
Group (*n*)[1]	IG (*n* = 441)	CG (*n* = 429)	IG (*n* = 405)	CG (*n* = 388)	IG (*n* = 361)	CG (*n* = 322)
Eating behavior[4]	9.15 (0.45)	9.25 (0.44)	−1.24 (0.31)	−0.87 (0.26)	−1.57 (0.35)	−1.08 (0.33)
AN-Subscale[4]	2.50 (0.16)	2.46 (0.16)	−0.28 (0.10)	−0.18 (0.10)	−0.35 (0.12)	−0.31 (0.12)
Body self esteem[5]	9.38 (0.21)	9.63 (0.20)	0.79 (0.13)[a]	0.05 (0.12)[a]	0.54 (0.17)	0.13 (0.15)
Knowledge[6]	9.59 (0.14)	9.16 (0.14)	0.89 (0,15)[b]	0.29 (0.13)[b]	1.12 (0.14)[c]	0.45 (0.15)[c]
Risk-group (*n*)[1]	IG (*n* = 161)	CG (*n* = 158)	IG (*n* = 142)	CG (*n* = 142)	IG (*n* = 122)	CG (*n* = 112)
Eating behavior[4]	18.74 (9.19)	18.70 (8.97)	−3.31 (8.29)	−1.99 (7.45)	−4.62 (8.15)	−3.64 (8.00)
AN-Subscale[4]	5.73 (3.11)	5.55 (3.25)	−0.86 (2.84)	−0.57 (2.50)	−1.10 (3.12)	−1.12 (2.92)
Body self esteem[5]	5.96 (3.77)	6.77 (4.08)	1.01 (3.28)[a]	−0.06 (2.78)[a]	0.91 (3.67)	0.33 (3.44)
Knowledge[6]	8.47 (3.05)	8.04 (2.99)	0.94 (3.32)[b]	0.14 (2.67)[b]	1.24 (3.06)[c]	0.41 (2.82)[c]
Non-risk group (*n*)[1]	IG (*n* = 289)	CG (*n* = 279)	IG (*n* = 263)	CG (*n* = 246)	IG (*n* = 239)	CG (*n* = 210)
Eating behavior[4]	3.68 (2.69)	3.96 (2.58)	−0.12 (4.28)	−0.22 (2.92)	−0.14 (4.75)	−0.25 (3.71)
AN-Subscale[4]	0.71 (1.32)	0.71 (1.30)	0.02 (1.46)	0.05 (1.28)	−0.05 (1.59)	−0.11 (1.34)
Body self esteem[5]	11.37 (3.32)	11.22 (3.31)	0.67 (2.25)[a]	0.11 (2.10)[a]	0.32 (2.70)	0.03 (2.13)
Knowledge[6]	10.22 (2.71)	9.79 (2.72)	0.86 (2.81)[b]	0.37 (2.42)[b]	1.05 (2.47)[c]	0.47 (2.82)[c]

Superscripts a, b and c show significant differences ($p < .05$); all statistical tests are based on an analysis of covariance.
AN = Anorexia nervosa.
[1] Based on the eating behavior measurement.
[2] Mean difference baseline-post intervention.
[3] Mean difference baseline-follow-up.
[4] EAT: scale range eating behavior = 0–78; scale range AN-subscale = 0–12; lower scores indicate a less conspicuous eating behavior.
[5] FBeK: scale range = 0–15; higher scores indicate a greater body self esteem.
[6] Scale range = 0–15; higher scores indicate more knowledge.

The fact that there was no change in eating behavior emphasizes the need for further action, such as revising the project's contents. Another step has already been done with the development and implementation of a co-educational program for girls and boys attending the 7th grade ('Torera': Results 2011), covering AN, bulimia nervosa, binge-eating disorder and obesity. It also seems necessary to accompany the primary preventive intervention with secondary preventive actions. For this purpose, a telephone hotline with information about nearby counseling and treatment services has been established (++49 3641 937747).

Our results yielded interesting epidemiological aspects. For example, the proportion of girls with a higher risk for developing an eating disorder was at 36% (10 or more points in the EAT = subclinical conspicuous eating behavior after Buddeberg-Fischer and Reed, 2001). This proportion is higher than the risk ratio in the KiGGS reference sample, too (23.5%, cf. Hoelling and Schlack, 2007). The occurring difference between our study and the KiGGS study may be explained by the use of different instruments for detecting conspicuous eating behavior (Scoff, 5 Items vs. EAT, 26 items).

Another interesting finding is the high number of underweight participants at baseline. This percentage is more than doubled in comparison to the percentage in this age group in the KiGGS reference sample. We suppose that this bias could be explained because our study based the BMI calculations on self-reported height and weight. Other studies (e.g. Field et al., 2007) show that girls especially underestimate their weight and overestimate their height, resulting in a lower BMI and an inflated number of underweight participants.

Study limitations and strengths

The selectivity of the sample could have biased the results. Working with TME meant sacrificing randomized group assignments (see Berger et al., 2008). For this reason, the motivation to participate could vary between IG- and CG-schools, but not among individual students. Because PriMa was applied within the regular curriculum, all students were required to participate. Furthermore, post-hoc parallelization was used to monitor possible confounding effects by type of school, students per class, and survey period. The teachers were asked by questionnaires about their implementation as well as about their usage of add-on exercises to further control for confounders. Still, we could not draw any conclusions, because a large number of differences made it difficult to stratify our data.

Good data quality could be assured because the drop-out rate was maximal at 23%. Regarding a generally lower adherence among young participants, this value is only marginally above the threshold of 20%, which is usually accepted in clinical trials (such studies often tend to have higher percentages of drop-outs; e.g. Martin et al., 2006).

The possible negative effects formulated in earlier studies (Carter et al., 1997) and often apprehended by parents could be refuted by improvements in the IG and absence of deterioration in the CG.

Compared to other programs (see Stice et al., 2007), the effects' impact is less evident. However, because of small sample sizes in all reviewed programs (Stice and Shaw, 2004), we assume that most of the programs evaluated until now have not undergone a widespread dissemination. This point underlines the value of the orientation on the standards of the Society of Prevention Research (SPR, Flay et al., 2005). These earlier studies are also often carried out by external professionals, who are only temporarily available. In addition, they often target older girls (15 years old) or pre-selected high-risk groups. As such, they are not among the primary prevention programs following Caplan (1964). On the other hand, they ignore the epidemiological findings on the so-called 'sub-clinical' conspicuous eating behavior (Buddeberg-Fischer and Reed, 2001; Grilo, 2006) and the fact that the peak of incidence of AN is at age 15. According to the summary findings of two well-known meta-analyses (Stice and Shaw, 2004; Stice et al., 2007), these aspects increase the effect sizes. Many programs indeed meet scientific criteria for success, but our experience shows it could be questioned if they meet preventive care needs and the expectations of cooperation partners such as schools, municipalities, and ministries.

Conclusion

Taking into account the positive results of the pilot study (Level-1, including process evaluation) and the successful implementation of the program in cooperation with the TME and the Heidelberg Prevention Center (www.h-p-z.de), PriMa is a girl-specific, school-based primary preventative method, and is therefore a low-threshold and socially fair intervention. With relatively little effort (9-hour project teaching) and low costs, the program is able to increase AN-specific knowledge and to enlarge, at least in short term, students' body self-esteem. Since negative effects of the program could be ruled out, PriMa suits as one building block for a concept of health promotion in the field of eating-related problems at school.

Conflict of interest statement
None.

Acknowledgments

We thank the financial support by the Federal Ministry for Education and Research (BMBF: Project no. 01EL0602, term 2006–2009).

Furthermore we thank our practice partner Thuringian Ministry of Education, especially Jutta Beinersdorf and Margit Luedecke for the didactic counseling and competent application of teacher trainings and all the girls for patiently filling out the questionnaires.

References

Becker, C.B., Smith, L., Ciao, A.C., 2005. Reducing eating disorder risk factors in sorority members: a randomized trial. Behav. Ther. 36, 245–254.
Berger, U., 2008. Essstörungen wirkungsvoll vorbeugen (Effective Prevention of Eating Disorders). Kohlhammer, Stuttgart.
Berger, U., Sowa, M., Bormann, B., Brix, C., Strauss, B., 2008. Primary prevention of eating disorders: characteristics of effective programs and how to bring them to broader dissemination. Eur. Eat. Disord. Rev. 16, 173–183.
Buddeberg-Fischer, B., Reed, V., 2001. Prevention of disturbed eating behavior: an intervention program in Swiss high school classes. Eat. Disord. 9 (2), 109–124.
Bulik, C.M., Sullivan, P.F., Tozzi, F., Furberg, H., Lichtenstein, P., Pedersen, N.L., 2006. Prevalence, heritability and prospective risk factors for anorexia nervosa. Arch. Gen. Psychiat. 63, 305–312.
Caplan, G., 1964. Principles of Preventive Psychiatry. Basic Books, New York.
Carter, J.C., Stewart, D.A., Dunn, V.J., Fairburn, C.G., 1997. Primary prevention of eating disorders: might it do more harm than good? Int. J. Eat. Disord. 22, 167–172.
Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences. Erlbaum, Hillsdale, NJ.
Eldridge, S.M., Ashby, D., Kerry, S., 2006. Sample size for cluster randomized trials: effect of coefficient of variation of cluster size and analysis method. Int. J. Epidemiol. 35, 1292–1300.
Erdfelder, E., Faul, F., Buchner, A., 1996. GPOWER: a general power analysis program. Behav. Res. Meth. Ins. C. 28, 1–11.
Fairburn, C.G., Cooper, Z., 1993. The Eating Disorder Examination. In: Fairburn, C.G., Wilson, G.T. (Eds.), Binge Eating: Nature, Assessment and Treatment. Guilford Press, New York, pp. 317–355.
Field, A.E., Aneja, P., Rosner, B., 2007. The validity of self-reported weight change among adolescents and young adults. Obesity 15, 2357–2364.
Flay, B.R., Biglan, A., Boruch, R.F., et al., 2005. Standards of evidence: criteria for efficacy, effectiveness and dissemination. Prev. Sci. 6, 151–175.
Garner, D.M., Garfinkel, P.E., 1979. The Eating Attitudes Test: an index of the symptoms of anorexia nervosa. Psychol. Med. 9, 273–279.
Gerlinghoff, M., Backmund, H., 1993. Magersucht und Bulimie. Verstehen und Bewältigen (Anorexia nervosa and bulimia nervosa: Understanding and Coping). Beltz, Weinheim.
Grilo, M.G., 2006. Eating and Weight Disorders. Hove: Psychology Press Taylor & Francis Group.
Hedges, L.V., Hedberg, E.C., 2007. Intraclass correlations for planning group-randomized trials in education. Educ. Eval. Policy An. 29, 60–87.
Hoek, H.W., van Hoeken, D., 2003. Review of the prevalence and incidence of eating disorders. Int. J. Eat. Disorder. 34 (4), 383–396.

Hoelling, H., Schlack, R., 2007. Eating disorders in children and adolescents. First results of the German health interview and examination survey for children and adolescents (KiGGS). Bundesgesundheitsbla. 50, 794–799.
Jacobi, C., Hayward, C., de Zwaan, M., Krämer, H.C., Agras, W.S., 2004. Coming to terms with risk factors for eating disorders: application of risk terminology and suggestions for a general taxonomy. Psychol. Bull. 130, 19–65.
Kleimann, M., Moessle, T., 2008. The logs of Eliza and other media stories. Behavioral and developmental effects of a school based media education program—Berlin longitudinal study media. Int. J. Behav. Dev. 32, 55–59.
Kromeyer-Hauschild, K., Wabitsch, M., Kunze, D., et al., 2001. Perzentile für den Body-mass- Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben (Percentiles of the body mass index (BMI) for children and adolescents in German samples). Monatsschr. Kinderh. 149, 807–818.
Levine, M.P., Smolak, L., 2006. The prevention of eating problems and eating disorders. Theory, Research, and Practice. Mahwah, NJ, Erlbaum.
Martin, J.L.R., Pérez, V., Sacristán, M., Rodríguez-Artalejo, F., Martínez, C., Álvarez, E., 2006. Meta-analysis of drop-out rates in randomised clinical trials, comparing typical and atypical antipsychotics in the treatment of schizophrenia. Eur. Psychiat. 21, 11–20.
Raich, R.M., Portell, M., Peláez-Fernández, M.A., 2009. Evaluation of a school-based programme of universal eating disorders prevention: is it more effective in girls at risk? Eur. Eat. Disord. Rev. 18, 49–57.
Raudenbush, S.W., Bryk, A.S., 2002. Hierarchical linear models. Applications and Data Analysis Methods. SAGE, Newburg Park, London, New Delhi.
Roehrig, M., Thompson, J.K., Brannick, M., van den Berg, P., 2005. Dissonance-based eating disorder prevention program: a preliminary dismantling investigation. Int. J. Eat. Disorder. 39 (1), 1–10.
Schwartze, D., Sowa, M., Bormann, B., Brix, C., Wick, K., Strauss, B. & Berger, U. (accepted). Efficacy of the 'Teenage Obesity Prevention Program' (TOPP) for the prevention of overweight and low physical activity in boys attending 6th grade. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.
Steinhausen, H.C., 2002. The outcome of anorexia nervosa in the 20th century. Am. J. Psychiat. 159 (6), 1284–1293.
Stewart-Brown, S., 2006. What is the evidence on school health promotion in improving health or preventing disease and, specifically, what is the effectiveness of the health promoting schools approach? Health Evidence Network report. WHO Regional Office for Europe, Copenhagen http://www.euro.who.int/document/e88185.pdf. accessed 16.08.2010.
Stice, E., Shaw, H., 2004. Eating disorder prevention programs: a meta-analytic review. Psychol. Bull. 130, 206–227.
Stice, E., Trost, A., Chase, A., 2003. Healthy weight control and dissonance-based eating disorder prevention programs: results from a controlled trial. Int. J. Eat. Disorder. 33, 10–21.
Stice, E., Shaw, H., Marti, C.N., 2007. A meta-analytic review of eating disorder prevention programs: encouraging findings. Annu. Rev. Clin. Psycho. 3, 207–231.
Strauss, B., Richter-Appelt, H., 1996. Fragebogen zur Beurteilung des eigenen Koerpers. Hogrefe-Verlag, Goettingen.
Tuschen-Caffier, B., 2005. Essstörungen: Klassifikation und Diagnostik (Eating disorders: Classification and Diagnostics). In: Baumann, U., Perrez, M. (Eds.), Huber, Bern.
Zipfel, S., Loewe, B., Reas, D.L., Deter, H.C., Herzog, W., 2000. Long-term prognosis in anorexia nervosa: lessons from a 21-year follow-up study. Lancet 355, 721–722.