Available online at www.sciencedirect.com

ScienceDirect

Comprehensive Psychiatry 55 (2014) 1773–1784

COMPREHENSIVE PSYCHIATRY

www.elsevier.com/locate/comppsych

Re-examining premature mortality in anorexia nervosa: A meta-analysis redux

Aparna Keshaviah[a], Katherine Edkins[a], Elizabeth R. Hastings[a,b], Meera Krishna[a], Debra L. Franko[a,c], David B. Herzog[a,d], Jennifer J. Thomas[a,d], Helen B. Murray[a], Kamryn T. Eddy[a,d,*]

[a]*Eating Disorders Clinical and Research Program, Massachusetts General Hospital, Boston, MA, USA*
[b]*The Amundsen Group, Burlington, MA, USA*
[c]*Northeastern University, Boston, MA, USA*
[d]*Department of Psychiatry, Harvard Medical School, Boston, MA, USA*

Abstract

Anorexia nervosa (AN) is reported to have the highest premature mortality of any psychiatric disorder, but recent meta-analyses may have inflated estimates. We sought to re-estimate mortality after methodological corrections and to identify predictors of mortality. We included 41 cohorts from 40 peer-reviewed studies published between 1966 and 2010. Methods included double data extraction, log-linear regression with an over-dispersed Poisson model, and all-cause and suicide-specific standardized mortality ratios (SMRs), with 95% Poisson confidence intervals. Participants with AN were 5.2 [3.7–7.5] times more likely to die prematurely from any cause, and 18.1 [11.5–28.7] times more likely to die by suicide than 15–34 year old females in the general population. Our estimates were 10% and 49% lower, respectively, than previously reported SMRs. Risk of premature mortality was highest in studies with older participants, although confounding by treatment was present. Gender, ascertainment, and diagnostic criteria also impacted risk.

1. Background

Anorexia nervosa (AN) is considered to have one of the highest premature mortality rates of any psychiatric disorder, with 0% to 25% of individuals with AN estimated to die prematurely [1,2]. In recent years, two meta-analyses investigating all-cause [2] and suicide-specific [3] mortality in AN were published. However, both suffered from inherent shortcomings that we believe yielded inflated estimates, including inconsistent outcome definitions, use of analytic methods ill-suited for rare outcomes, and data extraction errors. As a field, it is critical that, when advocating for insurance reimbursement, grant funding, and public recognition, we do not undermine ourselves by exaggerating mortality rates. With meta-analytic estimates of AN mortality increasingly disseminated to the public [4–6], we sought to examine the impact of such shortcomings. After applying a more rigorous approach to data synthesis, we re-estimated risk of premature mortality in AN. Below, we detail the missteps of the recent meta-analyses and potential remedies in order to elucidate issues that all researchers who conduct meta-analyses should consider.

Various metrics of mortality have been reported, including the crude mortality ratio (CMR = ratio of observed deaths to total subjects in a sample), mortality rate (observed deaths per person-years), and standardized mortality ratio (SMR = ratio of observed deaths to expected deaths in the general population). While the CMR is the most intuitive measure, and the mortality rate is well-suited for rare outcomes, the SMR is the only metric that contextualizes observed mortality with regard to expected mortality for the general population from which a cohort was drawn (typically matched by location, age, and gender), thus allowing for a standardized

Financial support: This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Conflicts of interest: The authors report no financial or other relationship relevant to the subject of this article.

* Corresponding author at: Eating Disorders Clinical and Research Program, Massachusetts General Hospital, 2 Longfellow Place, Suite 200, Boston, MA 02114, USA. Tel.: +1 617 724 1744; fax: +1 617 726 1595.
E-mail address: keddy@mgh.harvard.edu (K.T. Eddy).

http://dx.doi.org/10.1016/j.comppsych.2014.07.017
0010-440X/

comparison. The SMR is particularly important to utilize in meta-analyses that synthesize data from studies differing in era and geography, since it accounts for changes in expected mortality over time (e.g., due to aging of the population as a whole), and for differences in mortality by region (e.g., due to the presence or absence of centralized death registries or differential reporting standards). However, care must be taken when estimating the overall SMR with rare outcomes, as traditional meta-analytic software packages are not well-suited to handle records with zero outcome events.

In their analysis of the all-cause SMR, reported to be 5.9 with a 95% confidence interval (CI) of [4.2–8.3], Arcelus et al. [2] excluded studies with no deaths, which comprised 29% of their sample. Exclusion of these studies was necessary because the StatsDirect software used (StatsDirect Ltd, Cheshire, England) does not readily allow for inclusion of zero values in the outcome (email correspondence with Dr. A.J. Mitchell, May 2013). Preti et al. [3] overcame this limitation, which was also present in the Comprehensive Meta-Analysis software they used (Biostat, UK), by applying a continuity correction—adding a small constant to any cell count of zero (i.e., when no suicides occurred in a study); they reported a suicide-related rate ratio (which is strictly comparable to a suicide-related standardized mortality ratio, or SSR) and 95% CI of 35.4 [26.4–47.4]. However, use of continuity corrections can result in inflated estimates, as the authors themselves pointed out. Moreover, prior simulation research has shown that, depending on the size of the continuity correction utilized, the weighting of studies and thus the overall estimate of mortality can vary widely [7].

Although meta-analyses are typically presented as the definitive quantitative synthesis of a given topic of research, the results ultimately hinge on a series of critical decision points. Building upon the work of Arcelus et al. [2] and Preti et al. [3], we 1) analyzed the SMR without artificially altering the data through continuity correction; 2) employed more sophisticated analytic techniques that allowed us to include studies with zero deaths; and 3) re-extracted all source data in order to standardize the definition of the at-risk sample size for mortality (which was inconsistently calculated from one study to the next), remove duplicate studies, correct extraction errors, and collect additional data on study-level covariates for meta-regression. We hypothesized that our estimates of the SMR and SSR would be lower than those reported by Arcelus et al. [2] and Preti et al. [3] as a result of these adjustments.

2. Methods

2.1. Study selection and data extraction

Data from all studies included in the previous meta-analyses [8-47] were double extracted from the original texts by two assessors (A.K., M.K.), and any inconsistencies in entry or interpretation across assessors were reconciled. Combining the study selection criteria of Arcelus et al. [2] and Preti et al. [3], we included original peer-reviewed studies that were published in English between 1966 and 2010, reported a minimum of 1 year of follow-up, enrolled at least 15 participants, reported enough information to calculate the SMR, and reported intake eating disorder (ED) diagnosis (AN, bulimia nervosa, or binge eating disorder).

While re-extracting the data, we observed that the at-risk sample sizes of the studies meta-analyzed were inconsistently reported, depending on the focus of the original article. For example, some authors reported the total number of participants at entry, while others included only those consenting for additional follow-up, even though non-consenting participants were known to be alive, because mortality was not the primary focus of their research. These differences in sample size impact the calculation of mortality rates as well as the SMR. To standardize the calculation of mortality across studies, we defined the at-risk sample size for mortality to include all participants known to be alive, regardless of consent for follow-up, and excluded only those participants who were truly lost to follow-up and for whom mortality status was unknown.

Despite the efforts of Arcelus et al. [2] and Preti et al. [3] to exclude duplicate studies by including only the longest follow-up of a given study, we noticed that a number of studies included were based on the same data registries or cohorts (see Table 1). We eliminated all such duplicates from our dataset, retaining only the study with the longest follow-up and/or the most complete data reporting when duplicates existed. We also added a study by Remschmidt et al. [8] to cover data previously included in Hebebrand et al. [48], which we excluded due to duplication with three studies already included (Löwe et al. [9], Herpertz-Dahlmann et al. [10], and Steinhausen et al. [11]). In addition to eliminating duplicates, we corrected errors identified in the extracted datasets from both meta-analyses. Examples of such errors include extracting data for the entire population rather than for AN participants only, overlooking values that were reported in the manuscript but not in the abstract (resulting in existing data extracted as “not reported”), and simple data entry errors.

Due to insufficient extraction of covariate data, Arcelus et al. [2] were able to examine only age and BMI as predictors of the all-cause mortality rate via meta-regression. In our re-extraction, we gathered more data on covariates through a variety of strategies. For example, when covariate data were summarized separately by group (e.g., for males and females but not overall), we applied a weighted average based on group size to obtain an overall study-level value. When study-level covariate data were reported as a range, we used the midpoint, and when study-level covariates were not reported, we contacted the original authors to try to obtain the missing data. For two studies including a total of 3 participants with unknown causes of death (n = 1 reported by Hall et al. [12]; n = 2 reported by Møller-Madsen et al. [13]), we assumed that cause of death was not suicide. Finally, for some studies, we included covariate data that summarized all enrolled participants and not just those with

Table 1
Summary of studies included in the present or prior meta-analyses of mortality in anorexia nervosa.

Author	Year	Country	N[a]	Deaths	Suicides	SMR	SSR[b]	Reason for exclusion
Ben-Tovim et al[19]	1993	Australia	95	3	0	15.1[b]	0.0	
Birmingham et al[20]	1981	Canada	326	17	7	10.5	44.9	
Button et al[21]	1992	England	295	10	2	9.8	28.2	
Calvo Sagardoy et al[22]	1976	Spain	41	0	0	0.0	0.0	
Casper and Jabine[23]	1975	USA	75	4	1	10.4	29.5	
Coren and Hewitt[50]	1986	USA	Unk	571	8	NC	NC	Death certificate review (N unk)
Crisp et al[49]	1965	Scotland	63	8	4	4.7	35.9	⊆ Millar
Crisp et al[24]	1960	England	935	53	17	0.98	35.3	
Crow et al[25]	1985	USA	54	4	0	8.4	0.0	
Crow et al[26]	1979	USA	177	7	1	1.7	6.1	
Dancyger et al[51]	1979	USA	76	5	NR	10.4[b]	NC	Duplicates Eckert
Deter and Herzog[52]	1971	Germany	66	8	1	9.6	NC	Duplicates Lowe
Eckert et al[27]	1985	USA	76	5	0	12.8	0.0	
Emborg[53]	1977	Denmark	2763	231	46	22.3[b]	6.7	⊆ Moller-Madsen
Fichter et al[28]	1985	Germany	102	7	0	8.9	0.0	
Franko et al[54]	1987	USA	136	NR	4	NR	NR	Duplicates Keel
Hall et al[12]	1967	New Zealand	50	1	0	3.8	0.0	
Halvorsen et al[29]	1986	Norway	55	0	0	0.0	0.0	
Hebebrand et al[48]	1963	Germany	272	12	2	11.1[b]	16.3	Based on 5 studies already included
Herpertz-Dahlmann et al[10]	1985	Germany	39	0	0	0.0	0.0	
Herzog et al[55]	1987	USA	136	7	3	9.6	NC	Duplicates Keel
Joergensen[14]	1977	Denmark	81	7	NR	12.2	Unk	Excluded from SSR analysis (COD unk)
Keel et al[30]	1987	USA	136	10	4	11.6	56.9	
Korndörfer et al[31]	1935	USA	208	17	2	0.71	7.3	
Kreipe and Dukarm[33]	1979	USA	86	2	0	6.2	0.0	
Lee et al[32]	1984	Hong Kong	88	3	2	10.5	30.7	
Lindblad et al[56] Cohort 1	1977	Sweden	564	25	10	7.7	18.3	⊆ Papadopoulos
Lindblad et al[56] Cohort 2	1987	Sweden	554	7	2	2.9	3.9	⊆ Papadopoulos
Löwe et al[9]	1971	Germany	84	14	2	9.8	31.8	
Millar et al[34]	1965	Scotland	507	23	2	3.3	1.5	
Møller-Madsen et al[13]	1970	Denmark	853	50	18	9.1	69.9	
Morgan et al[35]	1973	England	78	1	1	4.4[b]	53.3	
Pagsberg and Wang[36]	1970	Denmark	30	2	0	6.5	0.0	
Papadopoulos et al[18]	1973	Sweden	6009	265	84	6.2	13.7	
Patton[37]	1971	England	332	11	6	6.0	56.2	
Ramsay et al[17] Detained	1983	England	79	10	0	54.6	0.0	
Ramsay et al[17] Voluntary	1983	England	78	2	1	11.1	73.6	
Ratnasuriya et al[38]	1959	England	41	7	3	15.4	82.5	
Remschmidt et al[8]	1952	Germany	101	3	0	6.5[b]	0.0	ADDED (to replace Hebebrand data)
Rosling et al[57]	1974	Sweden	157	21	7	11.7	42.1	⊆ Papadopoulos
Saccomani et al[39]	1976	Italy	85	0	0	0.0	0.0	
Santonastaso et al[40]	1970	Italy	46	2	0	15.2	0.0	
Signorini et al[41]	1994	Italy	138	3	1	9.7	43.1	
Steinhausen et al[11]	1979	Germany	55	4	1	15.9	35.5	
Strober et al[42]	1980	USA	95	0	0	0.0	0.0	
Sullivan et al[43]	1981	New Zealand	74	1	1	1.7[b]	16.0	
Tanaka et al[16]	1982	Japan	61	7	2	35.7[b]	46.6[b]	
Theander[44]	1931	Sweden	94	12	3	1.5	31.3	
Tolstrup et al[45]	1960	Denmark	151	9	6	4.8	87.4	
Touyz and Beumont[15]	1972	South Africa	34	2	1	0.99	748.8	Excluded from SSR analysis (outlier)
Van Son et al[46]	1985	Netherlands	76	1	1	6.9	44.7	
Wentz et al[47]	1985	Sweden	51	0	0	0.0	0.0	

Highlighted studies were excluded from our analyses; NR/NC = Not reported/calculated; Unk = Unknown; ⊆ stands for "is a subset of "; COD = causes of death; SMR/SSR = all-cause/suicide-specific standardized mortality ratio.

[a]Defined as all participants known to be alive, excluding those lost to follow-up and for whom mortality status was unknown.

[b]Calculated using data on expected deaths (except for Keel et al., 2003, who reported the SSR in their manuscript).

Table 2
Changes in SMR and SSR based on changes in data extraction and statistical analysis methods.

Prior meta-analysis Re-analyzed	Methodological changes made: Data extraction – Fixed errors[a]	Statistical methods – Used Poisson regression[b]	Study selection – Included additional studies that met criteria[c]	Software Used	SMR [95% CI]	SSR [95% CI][d]
Arcelus et al[2]				StatsDirect	5.9 [4.2 – 8.3]	
Arcelus et al[2]	X			StatsDirect	5.3 [3.5 – 8.0]	
Arcelus et al[2]		X		SAS	5.6 [4.0 – 8.0]	
Arcelus et al[2]	X	X		SAS	4.8 [3.2 – 7.3]	
Arcelus et al[2]	X		X	StatsDirect	6.9 [4.8 – 9.7]	
Arcelus et al[2]	X	X	X	SAS	5.2 [3.7 – 7.5]	
Preti et al[3]				CMA		35.4 [26.4 – 47.4]
Preti et al[3]	X			CMA		33.0 [22.7 – 48.0]
Preti et al[3]		X		SAS		26.0 [18.2 – 37.1]
Preti et al[3]	X	X		SAS		19.6 [12.3 – 31.4]
Preti et al[3]	X		X	CMA		32.8 [23.0 – 46.8]
Preti et al[3]	X	X	X	SAS		18.8 [11.9 – 29.6]

CI = Confidence Interval;
[a] Removed duplicate studies / cohorts (see Table 1 for details) and corrected mortality calculations by correcting the sample size considered at risk for all-cause mortality (SMR) and the years used to calculate the expected suicide rate (SSR);
[b] A log-linear regression model with an over-dispersed Poisson model for the error distribution was fit (over-dispersion captures random effects); the model allows for inclusion of studies with zero deaths (unlike Arcelus et al's approach) without the need for continuity correction (unlike Preti et al.'s approach), and is suitable for rare events;
[c] See Table 1 for details of additional studies included (includes studies reported by Preti et al. but not by Arcelus et al., and vice-versa);
[d] Results obtained using the Comprehensive Meta-analysis (CMA) software were based on the rate ratio (RR = observed divided by expected suicide rates), which is comparable to the SSR.

known mortality status, because the latter was not always available. All study-level covariates analyzed are listed in Table 2 (data available upon request). Categorical covariates with sparse data (i.e., <10% of studies) for some levels were collapsed based on similarity of the categories involved (e.g., based on similarity of diagnostic criteria).

In total, Arcelus et al. [2] included 35 studies describing mortality in AN, while Preti et al. [3] included 42 distinct AN cohorts from 40 studies, resulting in 51 distinct cohorts across both meta-analyses (26 studies were common to both). After excluding 11 cohorts due to duplication [48-57] and adding in the Remschmidt et al. [8] study, our meta-analysis included a total of 41 distinct cohorts from 40 studies. Analyses of suicide-specific mortality (SSR) excluded Joergensen [14] because they did not report cause of death, and Touyz and Beumont [15] because the calculated SSR of 748.8 was an extreme outlier (due to only 1 observed suicide and a very small expected number of suicides in the general population in South Africa)[1].

2.2. Statistical methods

When not reported in the original study, the SMR was calculated as the number of observed deaths divided by the number of expected deaths in the study, based on data double extracted by two of three assessors (A.K., K.E. or E.H.) from the World Health Organization Statistical Information System (WHOSIS) [58], the Office for National Statistics (ONS) (for England and Wales) [59,60], and the General Register Office for Scotland [61–64]. The expected number of deaths was derived by multiplying the study's sample size by the proportion dying (i.e., the CMR) in the general population. For each study, the CMR was calculated by summing the number of deaths (or suicides) that occurred in the general population during the years of study follow-up[2], and dividing this sum by the number alive at the beginning of the follow-up interval. When the follow-up years of a study were not specified, we used the year of study initiation through the year of initiation plus mean follow-up. When data were not available for the interval of interest, we selected the closest interval (of the same size) with data, prioritizing recent over older data. Like Preti et al. [3], we extracted data from the national registries for females aged 15–34, which represented the majority of participants enrolled.

Simulation research has shown that many commonly used statistical methods for meta-analysis (which have been implemented by traditional meta-analytic software packages) have serious deficiencies and may yield inappropriate results

[1] Preti et al (2011) incorrectly attributed the country for Touyz and Beumont (1984) as Australia.

[2] Preti et al (2011) used a fixed 5-year window ending with the publication year instead of the actual years of study follow-up.

when the outcome is rare and thus events are sparse [65,66]. To estimate the overall SMR, we generated a log-linear model, which regresses the log(SMR) against a constant value. After algebraic rearrangement, this results in a model where the log(observed deaths) is regressed against a constant plus the log(expected deaths) as the "offset" term. Estimation of the log(observed deaths) was conducted in PROC NLMIXED in SAS v9.3 (SAS Institute, Cary, NC) using an over-dispersed Poisson model for the error distribution (over-dispersion captures random effects). Poisson regression is well-suited to model counts of rare events, and allows for inclusion of studies with zero observed deaths [67]. To identify predictors of variability in AN mortality, we added study-level covariates to this Poisson regression model. Covariates were examined with and without outliers identified through visual inspection of plots of the log(SMR) against each covariate; the designation of a study as an outlier differed by covariate. Each covariate was tested univariately first, and those that were significant (with or without outliers) were then entered into a multivariate model. The final multivariate model was then obtained through stepwise selection ($p < 0.10$ to enter, $p < 0.05$ to stay) and excluded outliers. All analyses were conducted using SAS v9.3 (SAS Institute, Cary, NC).

3. Results

Our meta-analysis of 41 studies/cohorts included a total of 12,071 participants diagnosed with AN, who contributed 154,590 person-years of follow-up. A total of 579 participants died from any cause, with 169 having died by suicide. Mean age at study entry ranged from 14 to 27 years, mean years of ED at entry ranged from 0 to 8 years, mean percent of ideal body weight ranged from 64% to 79% at entry, and studies were largely comprised of females (males comprised between 0% and 18% of the populations studied).

3.1. Estimates of all-cause and suicide-specific mortality

Of the 41 studies/cohorts included in our meta-analysis, 39 had an SMR less than 16.0 (including 6 with an SMR = 0.0), and the 2 remaining studies had an SMR of 35.7 [16] and 54.6 [17]. The median and mean SMR was 6.5 and 8.5, respectively. Based on Poisson regression, we estimated an overall SMR [95% Poisson CI] of 5.2 [3.7–7.5]. To examine the robustness of this estimate, we ran sensitivity analyses that excluded one study at a time from the calculation and found that the overall SMR was quite stable, ranging from 4.9 to 5.7.

Among the 39 studies/cohorts included in our meta-analysis of suicide-specific mortality, 16 studies (41%) observed zero suicides (SSR = 0), and the SSR ranged from 1.5 to 87.4 in the remaining 23 studies. The median and mean SSR was 13.7 and 23.7, respectively. The overall SSR [95% Poisson CI] was 18.1 [11.5–28.7]. When we excluded one study at a time from the calculation in sensitivity analyses, the SSR ranged from 17.0 to 22.9.

To better understand the sources of difference in the estimates of mortality between our re-analysis and the previous meta-analyses, we generated follow-up analyses whereby we applied the analytic methods and software packages used by Arcelus et al. [2] and Preti et al. [3] to our re-extracted dataset, and conversely we applied our analytic methods to their original data extractions. Table 2 shows how the SMR and SSR change depending on the methodological changes applied. When we considered only those studies originally included by Arcelus et al. [2], but removed duplicate studies/cohorts and fixed data extraction errors, we estimated the overall SMR [95% Poisson CI] to be 5.3 [3.5–8.0]. When we additionally applied more sophisticated modeling techniques to these data to allow for inclusion of studies with 0 deaths, we obtained an estimate [95% CI] of 4.8 [3.2–7.3]. We hypothesized that our estimate of the overall SMR would be lower than that reported by Arcelus et al. [2] because we included studies with zero deaths, and also because many of the duplicate (and thus over-weighted) studies we removed had an SMR higher than the overall weighted average (Table 1). Table 2 shows that although removal of duplicates diminished the overall SMR, the effect was coincidentally almost canceled out by the addition of studies analyzed by Preti et al. [3] but not by Arcelus et al. [2], many of which had SMRs above average. For suicide-specific morality, when we considered only those studies originally included by Preti et al. [3], but removed duplicate studies/cohorts and fixed data extraction errors, we estimated the overall SSR [95% Poisson CI] to be 33.0 [22.7–48.0]. For suicide-specific mortality, when we additionally applied more sophisticated modeling techniques to these data to allow for inclusion of studies with 0 suicides without the need for continuity correction, we obtained an estimate [95% CI] of 19.6 [12.3–31.4].

3.2. Predictors of all-cause and suicide-specific mortality

Table 3 summarizes study-level covariates among studies with the highest and lowest SMRs (approximate median split), along with their predictive value in univariate regression. A number of covariates that were mean age at entry, mean years of ED was no longer significant. After step-wise model selection, mean age at entry, percent of males enrolled, and percent ascertainment significantly predicted variability in all-cause mortality. Higher SMRs were observed among studies that enrolled participants with an older mean age at entry [β(SE) = 0.11 (0.049), $t(39) = 2.22$, $p = 0.032$], enrolled fewer males [β(SE) = −0.12 (0.045), $t(39) = -2.71$, $p = 0.010$], and had lower percent ascertainment [β(SE) = −0.072 (0.034), $t(39) = -2.11$, $p = 0.041$]. However, we noted that age was highly correlated with treatment (point-biserial correlation = −0.47, $p = 0.002$), and that the mean age at entry was significantly higher in non-treatment (mean age =

Table 3
Predictors of the all-cause standardized mortality ratio (SMR).

	Lowest SMRs (<7.0, n=22)		Highest SMRs (>8.0, n=19)		Univariate regression results			
Study–level covariate	Mean	SD	Mean	SD	β Estimate	Std. Error	t–value (df)	p–value
Year of initiation	1970	15.0	1981	8.9	0.035	0.010	3.47 (40)	0.001*
Sample size	423	1265	145	188	–0.0051[a]	0.16[a]	–0.03 (40)	0.98
Follow–up years (mean)	12	7.1	10	4.4	–0.056	0.025	–2.24 (40)	0.03*
Person–years of follow–up (total)	5937	17115	1262	1458	–0.0027[a]	0.012[a]	–0.23 (40)	0.82
% with ≥1 co–morbid disorder at entry	52%	20.1%	44%	27.2%	–0.018	0.016	–1.16 (12)	0.27
% Males	5%	4.7%	2%	3.0%	–0.12	0.048	–2.56 (40)	0.01*
Age at entry (mean)	19	3.1	22	3.1	0.14	0.053	2.61 (40)	0.01**
Years of ED at entry (mean)	2	1.3	4	2.5	0.35	0.091	3.82 (26)	<0.001**
BMI at entry (mean)	16	2.0	15	1.5	–0.20	0.14	–1.45 (20)	0.16
% ideal body weight at entry (mean)	73%	3.7%	69%	5.6%	–0.059	0.048	–1.23 (10)	0.25
% Ascertainment	98%	6.5%	97%	5.6%	–0.33	2.86	–0.12 (40)	0.91[b]
	N	%	N	%				
Treatment given: Yes (vs. No)	10	45%	3	16%	–0.80	0.42	–1.93 (40)	0.06
Study type: Prospective (vs. Retro–)	11	50%	11	58%	0.19	0.35	0.55 (40)	0.59
Diagnostic criteria							F=0.48	0.63
ICD–10/DSM–IIIR/DSM–IV	13	59%	11	58%	ref.	–	–	–
Russell/Feighner/DSM–III/ICD–9	7	32%	4	21%	–0.027	0.42	–0.06 (40)	0.95
Ad–hoc/Unknown/ICD–8	2	9%	4	21%	0.44	0.48	0.92 (40)	0.36
Method of assessment							F=0.09	0.91
(Semi–) Structured Interview	11	50%	11	58%	ref.	–	–	–
Questionnaire/Unstructured	5	23%	2	11%	–0.19	0.51	–0.38 (40)	0.71
Record Review	6	27%	6	32%	–0.12	0.39	–0.31 (40)	0.76
Population type							F=1.22	0.31
Inpatient	8	36%	9	47%	ref.	–	–	–
Outpatient/Inpatient combination	12	55%	9	47%	–0.51	0.35	–1.46 (40)	0.15
Outpatient	2	9%	1	5%	–0.67	0.71	–0.94 (40)	0.35
Country							F=0.75	0.56
USA	4	18%	4	21%	ref.	–	–	–
England	3	14%	4	21%	0.69	0.54	1.28 (39)	0.21
Germany	2	9%	3	16%	0.68	0.61	1.10 (39)	0.28
Other European	10	45%	4	21%	0.095	0.48	0.20 (39)	0.85
Other	3	14%	4	21%	0.54	0.56	0.97 (39)	0.34

Highlighted covariates were significant multivariately. Std. = standard; df. = degrees of freedom; Ref. = referent group.
[a]Based on an increase of 1000 units in the covariate.
[b]$p < 0.10$ after excluding outliers.
*Significant at alpha = 0.05 level overall, but non-significant after excluding outliers.
**Significant at alpha = 0.05 level even after excluding outliers (if any).

21.7 years, SD = 3.2) versus treatment (mean age =18.2 years, SD = 3.1) studies. When treatment was added into the final model, neither age nor treatment was significant, indicating that the effect of age on AN mortality was confounded by treatment. Fig. 1(A–C) plots the log of the SMR (calculated using a continuity correction of 0.5, in order to graphically represent the log(SMR) when the SMR=0) against significant multivariate predictors (note that these figures only approximate the actual model that was generated, since Poisson regression does not apply a continuity correction).

Table 4 summarizes study-level covariates among studies with the highest and lowest SSRs (median split), as well as their predictive value in univariate regression. With the exception of mean follow-up time, all other covariates that were significant univariate predictors of SSR remained significant even after excluding a few outliers. Diagnostic criteria and mean age at entry were significant multivariate predictors, with higher SSR values observed in studies of participants with an older mean age at entry [β(SE) = 0.13 (0.060), $t(38) = 2.22$, $p = 0.033$] and in studies that used either Ad hoc/Unknown/ICD-8 diagnostic criteria [β(SE) =1.17 (0.41), $t(38) = 2.82$, $p = 0.0077$] or Russell/Feighner/DSM-III/ICD-9 diagnostic criteria [β(SE) = 0.85 (0.41), $t(38) = 2.05$, $p = 0.047$] as opposed to the ICD-

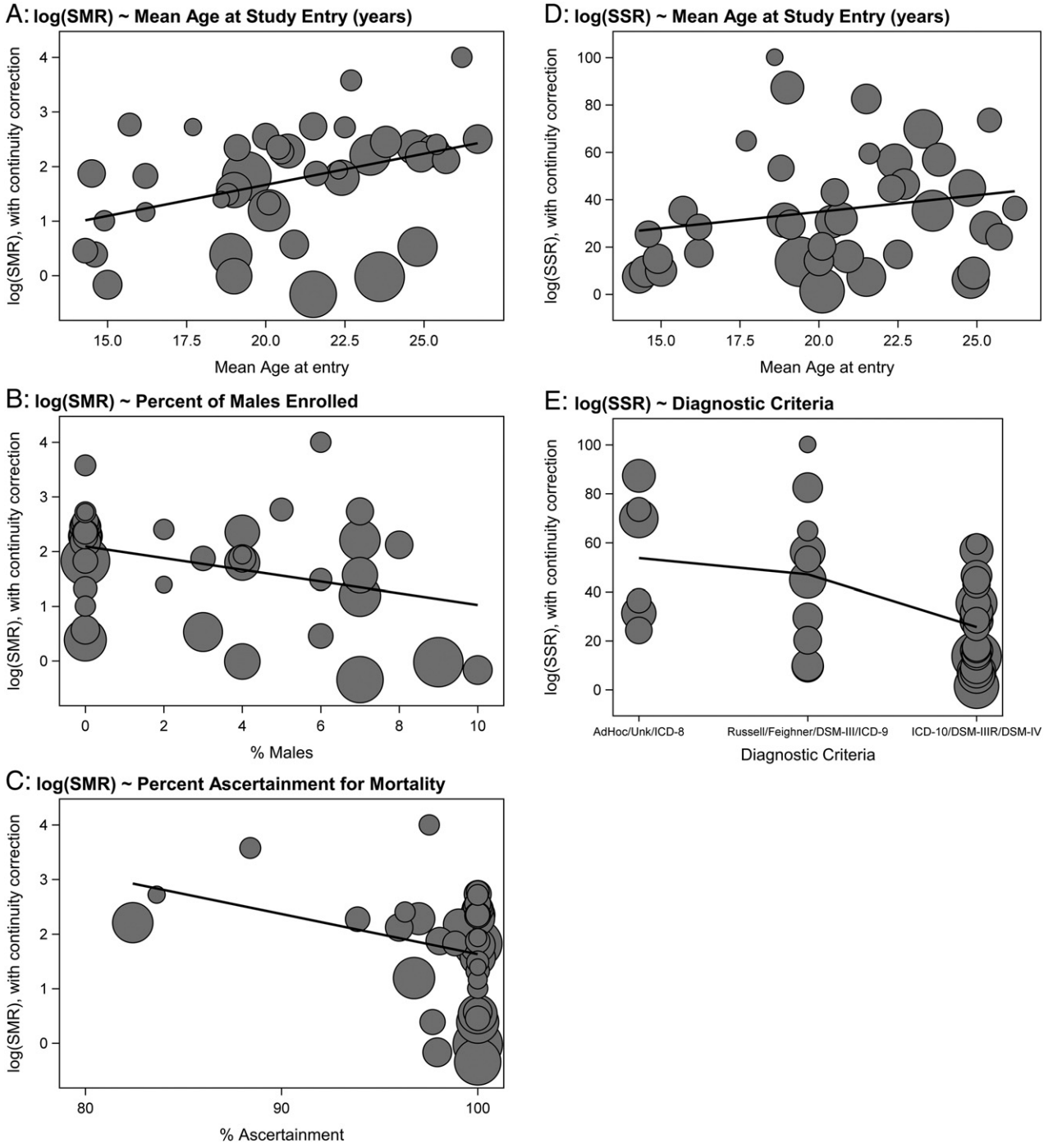

Fig. 1. Predictors of variability in the all-cause (SMR) and suicide-specific (SSR) standardized mortality ratio. Bubble plots show changes in all-cause (A–C) and suicide-specific (D–E) mortality by significant predictors (alpha = 0.10), with linear regression lines overlaid for continuous covariates (A–D), and lines connecting the means of categorical covariates (E). Bubbles are sized according to the log of the expected number of deaths/suicides (i.e., the "offset" term in the Poisson model), and outlying values of each covariate have been excluded. Each panel only approximates the Poisson model used, since Poisson regression does not apply a continuity correction, but continuity corrections (of 0.5 and 0.1, respectively) were required in on order to visually represent the log(SMR) and log (SSR) for studies with 0 deaths/suicides.

10/DSM-IIIR/DSM-IV [overall effect of diagnostic criteria: $F(2,38) = 4.77$, $p = 0.014$]. Although the percentage of individuals with at least one co-morbid disorder was also a significant predictor of the SSR univariately, we did not include this covariate in multivariate model selection because it restricted the effective sample size to 13 studies (due to lack of reporting on co-morbidity). Fig. 1(D–E) plots the log of the SSR (calculated using a continuity correction of 0.1) against these multivariate predictors (again, these figures only approximate the actual models generated).

Table 4
Predictors of the suicide-specific standardized mortality ratio (SSR).

	Lowest SSR (<8.0, n=19)		Highest SSR (>13.0, n=20)		Univariate regression results β	Std.	t–value	
Study–level covariate	Mean	SD	Mean	SD	Estimate	Error	(df)	p–value
Year of initiation	1976	13.8	1975	14.3	–0.0034	0.014	–0.24 (38)	0.81
Sample size	104	108	499	1321	–0.018[a]	0. 15[a]	–0.12 (38)	0.91
Follow–up time (mean years)	12	7.7	10	4.4	–0.091	0.024	–3.76 (38)	<0.001*
Person–years of follow–up (total)	1887	3939	5878	17706	–0.0069[a]	0.011[a]	–0.62 (38)	0.54
% with ≥1 co–morbid disorder at entry	61%	19.0%	38	21.7	–0.036	0.010	–3.54 (12)	0.004**
% Males	5%	5.1%	3%	3.2%	–0.014	0.056	–0.24 (38)	0.81
Age at entry (mean)	19	4.1	21	2.5	0.15	0.071	2.04 (38)	0.05**
Years of ED at entry (mean)	3	2.6	3	1.8	0.042	0.13	0.33 (25)	0.74[b]
BMI at entry (mean)	15	1.5	16	2.1	–0.069	0.18	–0.38 (20)	0.71
% of ideal body weight at entry	73%	3.7%	69%	6.0%	–0.040	0.065	–0.61(8)	0.56
% Ascertainment (mean)	98%	3.8%	98%	4.7%	–0.047	0.044	–1.07 (38)	0.29
	N	%	N	%				
Treatment given: Yes (vs. No)	9	47%	4	20%	–1.12	0.56	–1.99 (38)	0.05
Study type: Prospective (vs. Retro–)	13	68%	8	40%	–0.43	0.45	–0.96 (38)	0.34
Diagnostic criteria							F=4.16 (2,38)	0.02**
ICD–10/DSM–IIIR/DSM–IV	12	63%	11	55%	ref.	–	–	–
Russell/Feighner/DSM–III/ICD–9	2	11%	4	20%	1.24	0.46	2.72 (38)	0.01**
Ad–hoc/Unknown/ICD–8	5	26%	5	25%	0.75	0.43	1.73 (38)	0.09
Method of assessment							F=2.02 (2,38)	0.15
(Semi–) Structured Interview	12	63%	8	40%	ref.	–	–	–
Questionnaire/Unstructured	1	5%	6	30%	1.06	0.56	1.89 (38)	0.007**
Record Review	6	32%	6	30%	0.097	0.46	0.21 (38)	0.83
Population type							F=0.34 (2,38)	0.72
Inpatient	10	53%	7	35%	ref.	–	–	–
Outpatient/Inpatient combination	8	42%	11	55%	0.36	0.45	0.80 (38)	0.43
Outpatient	1	5%	2	10%	0.087	0.88	0.10 (38)	0.92
Country							F=1.53 (4,38)	0.21
USA	6	32%	2	10%	ref.	–	–	–
England	1	5%	6	30%	1.41	0.64	2.21 (38)	0.03**
Germany	3	16%	2	10%	0.074	0.86	0.09 (38)	0.93
Other European	7	37%	6	30%	0.56	0.60	0.93 (38)	0.36
Other	2	11%	4	20%	0.83	0.69	1.21 (38)	0.23

Highlighted covariates were significant multivariately. Std. = standard; Ref. = referent group.
[a]Based on an increase of 1000 units in the covariate.
[b]$p < 0.10$ after excluding outliers.
*Significant at alpha = 0.05 level overall, but non-significant after excluding outliers.
**Significant at alpha = 0.05 level even after excluding outliers (if any).

4. Discussion

After re-extracting all data and using statistical modeling techniques that allowed for inclusion of studies with zero deaths without the need for continuity corrections that artificially inflate estimates, we estimated that participants with AN are 5.2 times more likely to die prematurely from any cause, and 18.1 times more likely to die by suicide, when compared with 15–34 year old females in the population studied. Importantly, our estimate was 10% lower than the all-cause SMR reported by Arcelus et al. [2] and 49% lower than the suicide-specific rate ratio reported by Preti et al. [3]. When we restricted our analyses to include only those studies reported by the original authors (but removed duplicate studies/cohorts and corrected other extraction errors),

we obtained estimates of the SMR and SSR that were 17% and 45% lower, respectively, than previously reported. In other words, greater rigor applied during data extraction plus the use of more sophisticated statistical modeling techniques had a moderate to large impact on estimates of premature mortality in AN.

Our estimates of the overall SMR and SSR were not only lower than previously reported, but were also much less variable (based on tighter CIs) and robust to outliers. As an example, the extremely large sample size of Papadopoulos et al. [18], as shown in Table 1, would suggest that the study had undue influence on our results. On the contrary, we found the impact of this study to be negligible, with the SMR decreasing from 5.22 to 5.17, and the SSR increasing from 18.13 to 18.17, after excluding the study. The low variability in our estimates after excluding one study at a time lend further support to the stability of the analytic technique used, and we recommend that researchers analyzing rare outcome events such as mortality consider use of the over-dispersed Poisson regression model.

While lower than previously reported, our estimates of premature mortality in AN remain elevated and indicate significant risk that requires close monitoring in the clinical setting. Although the use of a more meticulous approach to the meta-analysis did not substantially alter the clinical implications in the case scenario we present, we feel that it is nonetheless important to report the most scientifically accurate estimates available, so that clinicians, patients, and their families have the most reliable information. Moreover, the use of more rigorous meta-analytic methods would be expected to have a greater impact on studies with outcomes that are less rare, i.e. when SMR estimates are more moderate in magnitude (e.g., if an overall SMR was estimated to be around 3, rather than in the 30s, a 50% reduction down to 1.5 would have greater implications). Furthermore, although multiple articles have asserted that AN has one of the highest premature mortality rates of any disorder, [2,4,5,68] after re-examining the source articles cited, we found the evidence to be ambiguous when comparisons were based on the SMR, which standardizes mortality against expected rates in a common population (and thus allows for a valid comparison). Based on meta-analyses of a small number of studies (generally under 10) differing by diagnosis, Harris and Barraclough [1] reported an all-cause SMR among males and females of all ages of 4.93 for AN, which was lower than that reported for bulimia (SMR = 9.38), but higher than SMRs for most other psychiatric disorders reported. By contrast, in a head-to-head comparison of mortality by diagnosis made within a single national cohort, Zilber et al. [69] reported SMRs ranging from 6.3 to 8.5 for 20–39 year old males and females diagnosed with schizophrenia, personality disorders, affective disorders, and other psychoses–all of which were higher than our estimate of AN mortality (SMR = 5.2). Thus, when comparisons are based on SMRs rather than on disparate mortality rates, elevated mortality in AN may not be out of scale with other psychiatric disorders.

Consistent with Arcelus et al. [2], and others [70], our analyses revealed higher risk of premature mortality in studies of older participants (based on mean age at entry) and those with relatively few males. However, by extracting additional covariate data, we were able to examine potential confounding. Although studies of participants with an older mean age at entry had higher SMRs than studies of younger participants, we found that the effect of age on mortality was confounded by treatment, as studies with older patients were also significantly less likely to provide/allow treatment as a part of the study design compared to studies with younger patients. Once treatment received was entered into the final multivariate Poisson regression model, age became non-significant, suggesting that the effect of age may in part be due to the effect of treatment, or that a third unknown variable may be driving the effects of both age and treatment. No significant correlations were observed to help explain why fewer deaths were observed in studies with a higher percentage of males enrolled (up to 10%, excluding outliers) than in those with few or no males enrolled, as seen in Fig. 1B. Further research is warranted to explore why premature mortality (based on the SMR) has been shown to be lower in males than in females with AN [49]. Our analysis also indicated that studies with better ascertainment had lower risk of premature mortality (note that although Table 3 shows only a 1 percentage difference in the average percent ascertainment between studies with the lowest and highest SMRs, ascertainment ranged from 82% to 100% and was a significant predictor at the alpha = 0.10 level after excluding a single outlying study with only 69% ascertainment). A plausible explanation for the effect of ascertainment on mortality is that studies with better ascertainment may have provided participants with a perception of connectedness or social support, and research has shown that positive experiences of social support are protective against mortality [71].

The highest risk of suicide-specific mortality was observed in studies enrolling participants with an older mean age at entry as well as those using ad hoc, unspecified, or older criteria to diagnose AN. Preti et al. [3] reported lower suicide rates (per 100 person-years) with increasing follow-up and with increasing year of publication. In our re-analyses, the effect of follow-up time became non-significant after removing two outlying studies, and we found no significant differences in SSR by year of study initiation, (which we believed better represented a study's time period than year of publication), likely because the SSR incorporates changes in background suicide rates over time.

There are some limitations to all meta-analyses that warrant discussion. Like our predecessors, during study selection we excluded studies with less than one year of follow-up, and thus our findings may not be fully generalizable. However, because follow-up time was not a significant predictor of SMR after the exclusion of two outliers, we do not believe that our generalizability was adversely affected. We also calculated the SMR using expected rates for females aged 15–34 in the general

population, which was representative of most patients enrolled in the studies included, but not all (on average, only 4% of patients [range: 0–18%] were male, and mean age at entry was 21 years [range: 14–27]). Additionally, because most studies included women aged 15–34, future mortality research is needed in younger, older, and male populations with AN. Another unavoidable limitation of all meta-regression analyses is ecological fallacy resulting from aggregation of data–i.e., that inference about participant-level characteristics is made based on relationships analyzed at the study level. However, when we compared our between-study findings with within-study findings, we were able to establish that the effects of age and gender on SMR when analyzed as study-level predictors in our meta-regression were in the same direction as when analyzed at the participant level within individual studies (few studies reported predictors of suicide, probably due to a lack of statistical power for a rare outcome). With respect to confounding, although meta-analyses may be based on randomized clinical trials, they themselves are observational in nature, and thus associations found through meta-regression cannot be taken as evidence for causality without further investigation.

5. Conclusions

Accurate mortality estimates for AN are crucial for clinical decision-making, legislative advocacy, and funding allocation. Meta-analyses are useful to synthesize data from multiple sources reported in a variety of ways, but for comparisons to be legitimate, standardization across studies is essential. During data extraction, researchers must consider whether data elements are consistently defined, and if not, should to re-calculate the outcome measure based on a common definition. Moreover, in meta-analyses of studies differing in location and era, the SMR is an important metric to consider in that it standardizes risk of premature mortality against expected rates and thus facilitates comparison across studies and disorders. Further, while traditional meta-analytic software packages are appealing because of their ease of use with simple outcomes (e.g., comparing group means), more sophisticated statistical models should be considered when analyzing rare outcomes, in order to account for zero outcome studies and to tighten confidence intervals. Finally, because meta-regression effects are particularly subject to confounding due to their observational nature and lack of individual level data, correlations between study-level covariates should be scrutinized to help elucidate underlying drivers of associations.

In the present meta-analysis, use of the aforementioned techniques resulted in lower estimates than previously reported. Moreover, review of the literature revealed that mortality in AN, while elevated, may be comparable to that of other psychiatric disorders when comparisons are based on the standardized mortality ratio. It is important that we do not undermine ourselves as a field by reporting and repeating inflated estimates. Indeed, AN need not have the highest mortality rate of any psychiatric disorder to be deserving of our clinical and research attention. As meta-analyses gain popularity among researchers and are increasingly cited by advocacy organizations, use of rigorous analytic methods is critical for valid and reliable results.

Acknowledgment

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

References

[1] Harris EC, Barraclough B. Suicide as an outcome for mental disorders. A meta-analysis. Br J Psychiatry 1997;170(3):205-28.
[2] Arcelus J, Mitchell AJ, Wales J, Neilson S. Mortality rates in patients with anorexia nervosa and other eating disorders. A meta-analysis of 36 studies. Arch Gen Psychiatry 2011;68(7):724-31.
[3] Preti A, Rocchi MB, Sisti D, Camboni MV, Miotto P. A comprehensive meta-analysis of the risk of suicide in eating disorders. Acta Psychiatr Scand 2011;124(1):6-17.
[4] Eating Disorders Coalition (EDC). Facts about eating disorders: what the research shows. http://www.eatingdisorderscoalition.org/documents/EatingDisorderFacts.pdf [Accessed June 28, 2013].
[5] Academy for Eating Disorders (AED). AED Report 2012: Eating Disorders, 2nd edition. http://www.aedweb.org/AM/Template.cfm?Section=Resources_for_Professionals&Template=/CM/ContentDisplay.cfm&ContentID=2984 [Accessed June 28, 2013].
[6] South Carolina Department of Mental Health. Eating disorder statistics. http://www.state.sc.us/dmh/anorexia/statistics.htm [Accessed June 28, 2013].
[7] Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. Stat Med 2004;23(9):1351-75.
[8] Remschmidt H, Wienand F, Wewetzer C. The long-term course of anorexia nervosa. In: Remschmidt H, & Schmidt MH, editors. Anorexia Nervosa. Toronto: Hogrefe & Huber; 1990. p. 127-36.
[9] Löwe B, Zipfel S, Buchholz C, et al. Long-term outcome of anorexia nervosa in a prospective 21-year follow-up study. Psychol Med 2001;31(5):881-90.
[10] Herpertz-Dahlmann B, Müller B, Herpertz S, et al. Prospective 10-year follow-up in adolescent anorexia nervosa–course, outcome, psychiatric comorbidity, and psychosocial adaptation. J Child Psychol Psychiatry 2001;42:603-12.
[11] Steinhausen H-C, Seidel R, Winkler Metzke C. Evaluation of treatment and intermediate and long-term outcome of adolescent eating disorders. Psychol Med 2000;30:1089-98.
[12] Hall A, Slim E, Hawker F, Salmond C. Anorexia nervosa: long-term outcome in 50 female patients. Br J Psychiatry 1984;145:407-13.
[13] Møller-Madsen S, Nystrup J, Nielsen S. Mortality in anorexia nervosa in Denmark during the period 1970–1987. Acta Psychiatr Scand 1996;94(6):454-9.
[14] Joergensen J. The epidemiology of eating disorders in Fyn County, Denmark, 1977–1986. Acta Psychiatr Scand 1992;85(1):30-4.
[15] Touyz SW, Beumont PJ. Anorexia nervosa. A follow-up investigation. Med J Aust 1984;141:219-22.
[16] Tanaka H, Kiriike N, Nagata T, Riku K. Outcome of severe anorexia nervosa patients receiving inpatient treatment in Japan: an 8-year follow-up study. Psychiatry Clin Neurosci 2001;55(4):389-96.
[17] Ramsay R, Ward A, Treasure J, Russell GFM. Compulsory treatment in anorexia nervosa: short-term benefit and long-term mortality. Br J Psychiatry 1999;175:147-53.

[18] Papadopoulos FC, Ekbom A, Brandt L, Ekselius L. Excess mortality, causes of death and prognostic factors in anorexia nervosa. Br J Psychiatry 2009;194(1):10-7.
[19] Ben-Tovim DI, Walker K, Gilchrist P, et al. Outcome in patients with eating disorders: a 5-year study. Lancet 2001;357(9264):1254-7.
[20] Birmingham CL, Su J, Hlynsky JA, Goldner EM, Gao M. The mortality rate from anorexia nervosa. Int J Eat Disord 2005;38(2):143-6.
[21] Button EJ, Chadalavada B, Palmer RL. Mortality and predictors of death in a cohort of patients presenting to an eating disorders service. Int J Eat Disord 2010;43:387-92.
[22] Calvo Sagardoy R, Fernandez Ashton A, Ayuso Mateos JL, Bayon Perez C, Santo-Domingo Carrasco J. Between 5 and 9 years follow-up in the treatment of anorexia nervosa. Psychother Psychosom 1989;52:133-9.
[23] Casper RC, Jabine LN. An eight-year follow-up: outcome from adolescent compared to adult onset anorexia nervosa. J Youth Adolesc 1996;25:499-517.
[24] Crisp A, Gowers S, Gowers N, et al. Death, survival and recovery in anorexia nervosa: a thirty five year study. Eur Eat Disord Rev 2006;14 (3):168-75.
[25] Crow S, Praus B, Thuras P. Mortality from eating disorders: a 5- to 10-year record linkage study. Int J Eat Disord 1999;26(1):97-101.
[26] Crow SJ, Peterson CB, Swanson SA, et al. Increased mortality in bulimia nervosa and other eating disorders. Am J Psychiatry 2009;166 (12):1342-6.
[27] Eckert ED, Halmi KA, Marchi P, Grove W, Crosby R. Ten-year follow-up of anorexia nervosa: clinical course and outcome. Psychol Med 1995;25(1):143-56.
[28] Fichter MM, Quadflieg N, Hedlund S. Twelve-year course and outcome predictors of anorexia nervosa. Int J Eat Disord 2006;39(2):87-100.
[29] Halvorsen I, Andersen A, Heyerdahl S. Good outcome of adolescent onset anorexia nervosa after systematic treatment: intermediate to long-term follow-up of a representative county-sample. Eur Child Adolesc Psychiatry 2004;13(5):295-306.
[30] Keel PK, Dorer DJ, Eddy KT, et al. Predictors of mortality in eating disorders. Arch Gen Psychiatry 2003;60(2):179-83.
[31] Korndörfer SR, Lucas AR, Suman VJ, et al. Long-term survival of patients with anorexia nervosa: a population-based study in Rochester, Minn. Mayo Clin Proc 2003;78(3):278-84.
[32] Lee S, Chan YYL, Hsu LKG. The intermediate-term outcome of Chinese patients with anorexia nervosa in Hong Kong. Am J Psychiatry 2003;160(5):967-72.
[33] Kreipe RE, Dukarm CP. Outcome of anorexia nervosa related to treatment utilizing an adolescent medicine approach. J Youth Adolesc 1996;25:483-97.
[34] Millar HR, Wardell F, Vyvyan JP, et al. Anorexia nervosa mortality in Northeast Scotland, 1965–1999. Am J Psychiatry 2005;162(4):165-72.
[35] Morgan HG, Purgold J, Welbourne J. Management and outcome in anorexia nervosa: a standardized prognostic study. Br J Psychiatry 1983:143282-7.
[36] Pagsberg AK, Wang AR. Epidemiology of anorexia nervosa and bulimia nervosa in Bornholm County, Denmark, 1970–1989. Acta Psychiatr Scand 1994;90(4):259-65.
[37] Patton GC. Mortality in eating disorders. Psychol Med 1988;18(4):947-51.
[38] Ratnasuriya RH, Eisler I, Szmukler GI, Russell GF. Anorexia nervosa: outcome and prognostic factors after 20 years. Br J Psychiatry 1991;158:495-502.
[39] Saccomani L, Savoini M, Cirrincione M, Vercellino F, Ravera G. Long-term outcome of children and adolescents with anorexia nervosa: study of comorbidity. J Psychosom Res 1998;44:565-71.
[40] Santonastaso P, Favaretto G, Canton G. Anorexia nervosa in Italy: clinical features and outcome in a long-term follow-up study. Psychopathology 1987;20:8-17.
[41] Signorini A, De Filippo E, Panico S, et al. Long-term mortality in anorexia nervosa: a report after an 8-year follow-up and a review of the most recent literature. Eur J Clin Nutr 2007;61(1):119-22.
[42] Strober M, Freeman R, Morrell W. The long-term course of severe anorexia nervosa in adolescents: survival analysis of recovery, relapse, and outcome predictors over 10–15 years in a prospective study. Int J Eat Disord 1997;22(4):339-60.
[43] Sullivan PF, Bulik CM, Fear JL, Pickering A. Outcome of anorexia nervosa: a case–control study. Am J Psychiatry 1998;155(7):939-46.
[44] Theander S. Anorexia nervosa: a psychiatric investigation of 94 female patients. Acta Psychiatr Scand Suppl 1970:2141-94.
[45] Tolstrup K, Brinch M, Isager T, et al. Long-term outcome of 151 cases of anorexia nervosa: the Copenhagen Anorexia Nervosa Follow-Up Study. Acta Psychiatr Scand 1985;71(4):380-7.
[46] van Son GE, van Hoeken D, van Furth EF, Donker GA, Hoek HW. Course and outcome of eating disorders in a primary care-based cohort. Int J Eat Disord 2010;43:130-8.
[47] Wentz E, Gillberg IC, Anckarsäter H, Gillberg C, Råstam M. Adolescent-onset anorexia nervosa: 18-year outcome. Br J Psychiatry 2009;194(2):168-74.
[48] Hebebrand J, Himmelmann GW, Herzog W, et al. Prediction of low body weight at long-term follow-up in acute anorexia nervosa by low body weight at referral. Am J Psychiatry 1997;154(4):566-9.
[49] Crisp AH, Callender JS, Halek C, Hsu LK. Long-term mortality in anorexia nervosa: a 20-year follow-up of the St George's and Aberdeen cohorts. Br J Psychiatry 1992:161104-7.
[50] Coren S, Hewitt PL. Is anorexia nervosa associated with elevated rates of suicide? Am J Public Health 1998;8:1206-7.
[51] Dancyger IF, Sunday SR, Eckert ED, Halmi KA. A comparative analysis of Minnesota Multiphasic Personality Inventory profiles of anorexia nervosa at hospital admission, discharge, and 10-year follow-up. Compr Psychiatry 1997;38(3):185-91.
[52] Deter HC, Herzog W. Anorexia nervosa in a long-term perspective: results of the Heidelberg–Mannheim Study. Psychosom Med 1994;56(1):20-7.
[53] Emborg C. Mortality and causes of death in eating disorders in Denmark 1970–1993: a case register study. Int J Eat Disord 1999;25:243-51.
[54] Franko DL, Keel PK, Dorer DJ, et al. What predicts suicide attempts in women with eating disorders? Psychol Med 2004;34(5):843-53.
[55] Herzog DB, Greenwood DN, Dorer DJ, et al. Mortality in eating disorders: a descriptive study. Int J Eat Disord 2000;28(1):20-6.
[56] Lindblad F, Lindberg L, Hjern A. Improved survival in adolescent patients with anorexia nervosa: a comparison of two Swedish national cohorts of female inpatients. Am J Psychiatry 2006;163(8):1433-5.
[57] Rosling AM, Sparén P, Norring C, von Knorring AL. Mortality of eating disorders: a follow-up study of treatment in a specialist unit 1974–2000. Int J Eat Disord 2011;44(4):304-10.
[58] World Health Organization Statistical Information System (WHOSIS). WHO Mortality Database. http://apps.who.int/healthinfo/statistics/mortality/whodpms/ [Accessed August 13, 2012].
[59] Office for National Statistics (ONS). Death registrations by sex and single year of age, 1961 to 2010, England and Wales. http://www.ons.gov.uk/ons/datasets-and-tables/index.html?newquery=mortality+death+vital&newoffset=50&pageSize=50&content-type=Reference+table&content-type=Dataset&content-type-orig=%22Dataset%22+OR+content-type_original%3A%22Reference+table%22&sortBy=none&sortDirection=none&applyFilters=true [Accessed September 25, 2012].
[60] Office for National Statistics (ONS). Suicide rates in the United Kingdom and in England and Wales, 1991 to 2010. http://www.ons.gov.uk/ons/taxonomy/index.html?nscl=Suicide+Rates#tab-data-tables [Accessed September 25, 2012].
[61] General Register Office for Scotland. Mid-year population estimates; Scotland by quinary age and sex: 1911 to 2011. http://www.gro-scotland.gov.uk/statistics/theme/population/estimates/mid-year/time-series.html [Accessed September 27, 2012].
[62] General Register Office for Scotland. Deaths, by sex and age groups, Scotland, 1901 to 2011. http://www.gro-scotland.gov.uk/statistics/theme/vital-events/deaths/time-series.html [Accessed September 27, 2012].
[63] General Register Office for Scotland. Table 1: Deaths for which the underlying cause was classified as 'intentional self-harm' or 'event of undetermined intent' by sex and by type of cause: registered in Scotland, 1974 to 2011. http://www.gro-scotland.gov.uk/statistics/

theme/vital-events/deaths/suicides/tables-and-chart.html [Accessed September 27, 2012].
[64] General Register Office for Scotland. Table 3: deaths for which the underlying cause was classified as 'intentional self-harm' or 'event of undetermined intent' by age-group: registered in Scotland, 1974 to 2011. http://www.gro-scotland.gov.uk/statistics/theme/vital-events/deaths/suicides/tables-and-chart.html [Accessed September 27, 2012].
[65] Brockwell SE, Gordon IR. A comparison of statistical methods for meta-analysis. Stat Med 2001;20:825-40.
[66] Bradburn MJ, Deeks JJ, Berlin JA, Russell LA. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. Stat Med 2007;26(1):53-77.
[67] Frome EL, Checkoway H. Use of Poisson regression models in estimating incidence rates and ratios. Am J Epidemiol 1985;121(2):309-23.
[68] Sullivan PF. Mortality in anorexia nervosa. Am J Psychiatry 1995;152(7):1073-4.
[69] Zilber N, Schufman N, Lernen Y. Mortality among psychiatric patients—the groups at risk. Acta Psychiatr Scand 1989;79(3):248-56.
[70] Haus C, Caille A, Godart N, et al. Factors predictive of ten-year mortality in severe anorexia nervosa patients. Acta Psychiatr Scand 2010;123(1):62-70.
[71] Croezen S, Haveman-Nies A, Picavet HS, et al. Positive and negative experiences of social support and long-term mortality among middle-aged Dutch people. Am J Epidemiol 2010;172(2):173-9.