



# Bariatric Outcomes Longitudinal Database (BOLD) Suggests Excess Weight Loss and Excess BMI Loss To Be Inappropriate Outcome Measures, Demonstrating Better Alternatives

Arnold van de Laar

Published online: 25 August 2012  
© Springer Science+Business Media, LLC 2012

**Abstract** Excess weight loss (%EWL) results show significant variation by initial body mass index (BMI): the heavier the patient, the lower the result. It is unclear whether this effect originates from the arithmetic construction of this outcome measure or from any true difference in bariatric impact on heavier and lighter patients. Outcome measures generating inappropriate variation would actually be unsuited for bariatric reports with possible implications on existing bariatric evidence. Nadir weight loss results after laparoscopic gastric bypass in 7,212 women from the Bariatric Outcomes Longitudinal Database (BOLD) are calculated for %EWL and 26 different relative measures using the formula  $100 \times (\text{initial BMI} - \text{nadir BMI}) / (\text{initial BMI} - a)$ , with reference point  $0 \leq a \leq 25$ . Variations by initial BMI and deviations in results using each relative measure are compared. Mean initial BMI,  $47.5 \text{ kg/m}^2$ . Mean nadir BMI,  $28.9 \text{ kg/m}^2$ . Mean nadir results of the lighter (A) and heavier (B) halves ( $2 \times n=3, 606$ ) by reference point,  $a=25$  (excess BMI loss (%EBMIL)),  $95.3 \%$  (A)– $77.2 \%$  (B);  $a=\text{ideal weight}$  (%EWL),  $81.7 \%$  (A)– $70.4 \%$  (B);  $a=10$ ,  $49.6 \%$  (A)– $49.7 \%$  (B);  $a=0$  (total weight loss (%TWL)),  $37.7 \%$  (A)– $40.2 \%$  (B). Variation by initial BMI is significant using all relative measures, except those with  $a=9$  ( $p=0.396$ ) and  $a=10$  ( $p=0.504$ ). The smallest variation coefficient is  $21.5 \%$  at  $8 \leq a \leq 14$ . Gastric bypass works equally effective for all female patients, reducing the part of body mass above  $10 \text{ kg/m}^2$  by an average of  $49.6 \%$ , regardless of their initial BMI. In contrast, %EWL and %EBMIL generate variation by initial BMI, providing lighter patients inappropriately with better results, potentially interfering with the

significance of bariatric evidence. These two most widespread used outcome measures in bariatric surgery are therefore actually unsuited for comparing patients or groups. This BOLD data makes a strong argument for abandoning %EWL and %EBMIL altogether and reporting bariatric relative outcome as %TWL only.

**Keywords** BOLD · Database · Excess weight loss · EWL · EBMIL · Relative outcome measures · Bariatric surgery · TWL · LRYGB · Weight loss · Weight loss results

## Introduction

Bariatric weight loss results can be expressed with absolute measures like kilograms and body mass index (BMI), or with relative measures. A 2005 review of the literature showed the simple relative measure total weight loss (%TWL) to be used both in surgical and nonsurgical reports, but the two relative measures excess weight loss (%EWL) and excess BMI loss (%EBMIL) to be favored by surgeons only [1]. %EWL and %EBMIL presently are the two most widespread used outcome measures in bariatric surgery, serving two purposes. First, they implicate a goal: ideal weight for %EWL and  $25 \text{ kg/m}^2$  for %EBMIL. In that sense %EWL and %EBMIL values express to what extent these goals are reached. The 25%EWL and 50%EWL marks (often mistaken for the 1981 Reinhold criteria [2]) are frequently used as criteria for failure and success of treatment. Next, %EWL and %EBMIL are used to compare weight loss, not to a preset goal, but to the results of other patients. In that sense they have become just outcome measures. Both applications of these measures are essentially different. %EWL and %EBMIL serve their original purpose (implicating a goal) as long as there is consensus among surgeons that ideal weight or  $25 \text{ kg/m}^2$  are the right

A. van de Laar (✉)  
Bariatric Surgery, Slotervaartziekenhuis,  
Louwesweg 6,  
1066EC Amsterdam, The Netherlands  
e-mail: arnoldvandelaar@usa.net

goals for bariatric surgery. This does not however make them suited automatically for their derived use as simple outcome measures. In fact, several authors reporting on different bariatric procedures observed that outcome expressed in %EWL varies by initial BMI: the heavier the patient, the lower the %EWL. This was noticed first for laparoscopic gastric bypass in 2000 by Higa et al. [3], for gastric banding in 2003 by Biertho et al. [4] and for biliopancreatic diversion in 2004 by Biron et al. [5]. These remarks however did not draw much attention. Recently, this finding was complemented by van de Laar et al. showing that the variation by initial BMI in %EWL results of a sample of 168 laparoscopic gastric bypass patients not only increased using %EBMIL but disappeared altogether using different relative measures [6]. This suggests that variation by initial BMI might not be intrinsic to bariatric outcome, but in fact caused by using specific relative measures, including %EWL and %EBMIL. These outcome measures might therefore be less suited for comparing patients with different initial BMI. However, the sample used by van de Laar et al. and the scarce reports on this variation in bariatric literature are not probative for this conclusion to be generally valid. A few questions remain. First, does variation by initial BMI always occur if %EWL and %EBMIL are used? Second, does it always disappear using other outcome measures? Third, is it relevant enough? If it were true that %EWL and %EBMIL generate variations in results that are both avoidable and relevant, than all existing evidence based on outcome expressed only as %EWL or %EBMIL could be affected and should be scrutinized.

The purpose of the present study is to test the first two questions by using the same arithmetic methods as described by van de Laar et al. in the largest series available, with data from the Bariatric Outcomes Longitudinal Database (BOLD). The relevance of these findings, the third question, is merely investigated.

In general, relative weight loss measures express weight loss relative to the initial BMI and to a specified reference body mass. In the case of %EWL, this reference is ideal weight, reflecting a patient's ideal body mass considering weight, length, gender, and body frame according to the 1983 Metropolitan Life Insurance Company Tables for ideal weight [7]. In the case of %EBMIL, this reference is the body mass with Quetelet index of 25 kg/m<sup>2</sup>, considering only weight and length. %TWL has no reference body mass; its reference point is 0. Many hypothetical relative weight loss measures can be constructed using different reference points. Using some might lead to more or less variation by initial BMI than using others.

## Materials and Methods

The Institutional Review Board approved the present study. The September 2011 BOLD database cut is searched for all

female patients that underwent a primary fully laparoscopic Roux-en-Y gastric bypass with minimum follow-up of 2 years. BMI at first preoperative visit (initial BMI) and age at time of surgery are determined. The nadir BMI is defined as lowest reported BMI postsurgery. Each absolute nadir weight loss result is transformed into 27 different relative nadir weight loss results. They are calculated with the formula:  $100 \% \times (\text{initial BMI} - \text{nadir BMI}) / (\text{initial BMI} - a)$  in which reference point “*a*” is any whole number ranging from 0 to 25 kg/m<sup>2</sup>, resulting in 26 datasets. An extra set is calculated for %EWL using mean values for medium body frame from the Metropolitan Life Insurance Company Tables for ideal weight. To demonstrate variation by initial BMI, the whole group is divided by initial BMI into halves and into quarters. Statistical significance of any difference in results between these subgroups is determined with the Mann–Whitney *U* test for each of the 27 relative outcome measures separately (considering a two-tailed  $p < 0.05$  significance). To further substantiate variation, the deviation of the results is determined for each of these 27 data sets. For this purpose, variation coefficients ( $VC = 100 \% \times \text{standard deviation} / \text{mean}$ ) are used instead of standard deviation because these sets of ratio variables are expected to have widely different means. It is important to note that in all of these operations, different calculations of the same bariatric outcome results are compared, not different bariatric outcome results mutually.

## Results

The September 2011 BOLD database cut yielded 7,340 female patients that underwent a primary fully laparoscopic Roux-en-Y gastric bypass with minimum follow-up of 2 years. One hundred twenty-eight patients were excluded because of a  $\geq 2.54$  cm (1 in. or more) difference between the reported preoperative and postoperative body length. Results of the remaining 7,212 patients are analyzed. Their mean characteristics and mean nadir weight loss results are presented in Table 1, expressed in different outcome measures and grouped by initial BMI in four quarters and two halves. Differences in initial BMI, age and body length between both halves (subgroups A and B) are significant ( $p < 0.05$ ). Median nadir results of these subgroups A (lighter) and B (heavier) are presented in Fig. 1, expressed in relative outcome measures with different reference points. They are, for  $a = 25$  (%EBMIL)—A 95.5 % and B 76.9 %; for  $a = \text{ideal weight}$  (%EWL)—A 82.1 % and B 70.3 %; for  $a = 10$ —A 50.3 % and B 49.6 %; and for  $a = 0$  (%TWL)—A 38.1 % and B 40.1 %, respectively.

### Significance of Variation by Initial BMI

The difference in relative results between subgroups A and B is significant ( $p < 0.05$ ) using %EWL ( $p < 0.0001$ ) and

**Table 1** Patient characteristics and nadir weight loss results expressed in different outcome measures of 7,212 women after primary fully laparoscopic gastric bypass, grouped by initial BMI in four quarters (Q1–4) and two halves (H1–2)

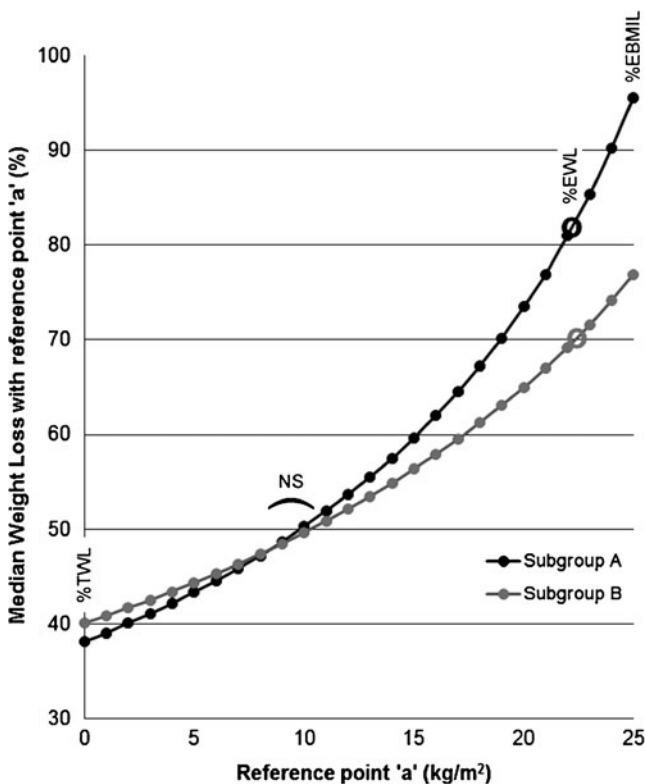
Group characteristics			Mean patient characteristics			Mean nadir weight loss					
Group	BMI range (kg/m <sup>2</sup> )	Size (n)	Age (year)	Body length (cm)	Initial BMI (kg/m <sup>2</sup> )	Nadir BMI (kg/m <sup>2</sup> )	Days to nadir	%TWL (a=0)	– (a=10 kg/m <sup>2</sup> )	%EWL (a=IW)	%EBMIL (a=25 kg/m <sup>2</sup> )
Q1	24.7–42.0	1,803	48.4	164.0	39.4	25.0	562	36.5	49.0	84.6	100.4
Q2	42.0–45.9	1,803	46.5	163.8	43.9	26.9	590	38.8	50.2	78.9	90.2
Q3	45.9–51.7	1,803	46.2	163.5	48.5	29.3	606	39.7	50.0	73.6	82.0
Q4	51.7–121.6	1,803	45.8	162.8	58.0	34.3	643	40.8	49.4	66.9	72.4
H1 (A)	24.7–45.9	3,606	47.5	163.9	41.7	25.9	576	37.7	49.6	81.7	95.3
H2 (B)	45.9–121.6	3,606	46.0	163.1	53.2	31.8	624	40.2	49.7	70.4	77.2
Total	24.7–121.6	7,212	47.5	163.5	47.5	28.9	600	38.9	49.6	76.1	86.2

*a* reference point of the relative measure, %TWL percent total weight loss (*a*=0), %EBMIL percent excess body mass index (BMI) loss (*a*=25 kg/m<sup>2</sup>), %EWL percent excess weight loss (*a* ideal weight (IW))

relative measures with *a* ≥ 11, including %EBMIL (*p* < 0.0001). This variation has an inversed correlation to the absolute weight loss: heavier patients show higher absolute weight loss but lower relative weight loss. There is no

significant variation by initial BMI using relative measures with *a*=9 (*p*=0.396) and *a*=10 (*p*=0.504). With these two relative measures, heavier and lighter patients show similar relative weight loss results.

There is a significant variation by initial BMI (*p* < 0.05) using relative measures with *a* ≤ 8, including %TWL (*p* < 0.0001). This variation is not inversed: heavier patients show both higher absolute weight loss and higher relative weight loss.



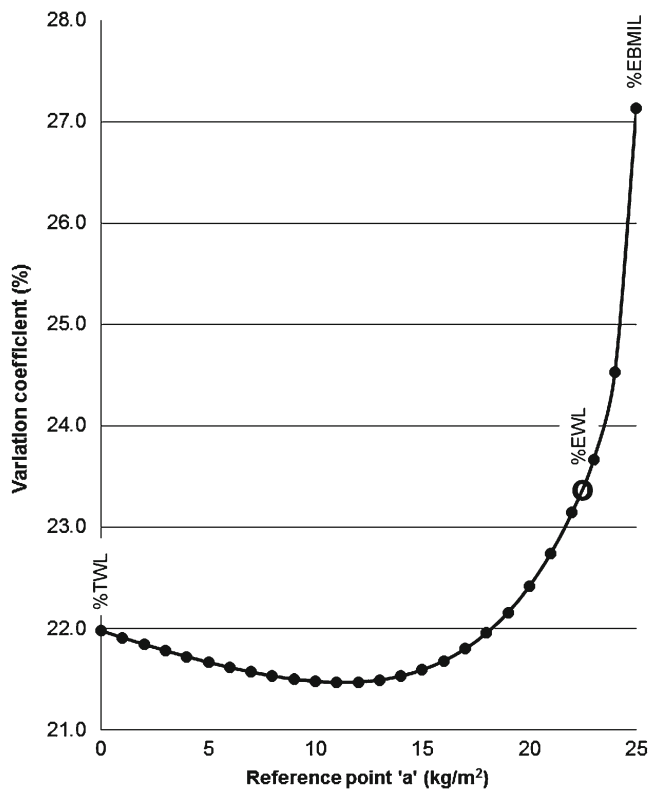
**Fig. 1** Variation by initial BMI. Medians of groups A (≤median initial BMI) and B (≥median initial BMI) of the same data expressed in different relative weight loss measures (in percent) according to the reference point “*a*” (in kilograms per square meter). %TWL percent total weight loss (*a*=0), %EBMIL percent excess body mass index loss (*a*=25), NS difference not significant (*p* > 0.05). The result for percent excess weight loss (%EWL) is projected on the curves (circles)

**Influence of Initial BMI on the Deviation of Results**

Variation coefficients for results expressed in different relative measures are presented in Fig. 2. The smallest variation coefficient is 21.5 % and is found at 8 ≤ *a* ≤ 14, the largest at *a* = 25 (%EBMIL).

**Discussion**

The significant (*p* < 0.0001) variation by initial BMI of weight loss results expressed in %EBMIL and %EWL in 7,212 patients from the BOLD database confirms previous findings that heavier patients show smaller %EBMIL and %EWL results than lighter patients. At first sight, it seems obvious that outcome measures based on excess weight (%EWL) or excess BMI (%EBMIL) are responsible for at least part of this effect. After all, heavier bariatric patients, having more absolute excess weight, need to lose more kilograms or BMI points than lighter patients in order to reach the same relative excess weight loss. It is however remarkable that even in this large series this variation disappears altogether when expressing the same results in a different relative measure (*p* = 0.504). This means that %EWL and %EBMIL are actually responsible not for part



**Fig. 2** Variation in deviation. Variation coefficients (in percent) as a measure of deviation of the same results expressed in different relative weight loss measures according to the reference point “a” (in kilograms per square meter). %TWL percent total weight loss ( $a=0$ ), %EBMIL percent excess body mass index loss ( $a=25$ ). The variation coefficient for percent excess weight loss (%EWL) is projected on the curve (circle)

of the effect but for all variation by initial BMI: heavier patients show smaller %EBMIL and %EWL results than lighter patients for the one reason that these specific measures are used. In other words, variation by initial BMI is not intrinsic to bariatric surgery, but in fact caused by using wrong outcome measures. It is unfortunate that the two most widespread used outcome measures in bariatric surgery, %EBMIL and %EWL turn out to be those wrong measures. The findings are supported by the clear differences in deviation using different relative measures to the detriment of %EWL and %EBMIL.

However, although all subjects in both subgroups A and B are female and underwent the same bariatric procedure, factors like ethnicity, eating habits or comorbidities were not taken into account in this study. Furthermore, the differences in age and body length between both subgroups might be very small (1.5 years in mean age and 0.8 cm in mean body length), but they are statistically significant. But, although these shortcomings could attribute to the variations seen in this study, they cannot explain the resolute disappearance of variation using different weight loss measures, which actually is the most important finding of this study.

Before concluding %EWL and %EBMIL to be unfit for comparing results of different patients altogether, the relevance of the variation these measures generate should be assessed. Relevance should not so much be found in the size of the variation itself, although the mean difference between the two outer quarters for %EBMIL is no fewer than 28 % (100 %EBMIL if initial BMI is <42 versus 72 %EBMIL if initial BMI is >52), but in the impact this redundant variation can have on the significance of conclusions. Conclusions in bariatric outcome studies are often based on significance of any difference (or absence of any difference) between two groups or within one. A difference in initial BMI could make a true difference in outcome look less significant than it should if expressed in %EWL or %EBMIL. For example, if there is no significant difference in %EWL results between a group of sleeve gastrectomy patients and a group of gastric bypass patients, it is still wrong to conclude that both bariatric procedures are equally effective if the sleeve patients initially were a bit lighter than the others. Same %EWL values in two patients with different initial BMI should in fact be interpreted as two different weight loss results, while two different %EWL values in two patients with different initial BMI could well be the expression of a similar weight loss. Mistakes could be made by neglecting these effects when omitting absolute results to accompany the relative outcome. This warning is not new. In 2005 Dixon et al. recommended presenting outcome in absolute terms in all cases to be a minimal reporting requirement [1]. Eleven years earlier, the committee on standards for reporting results of the American Society for Bariatric Surgery stated similar recommendations [8]. The relevance of %EWL-caused variation was illustrated in a recent systematic review examining 62 studies reporting on the potential association between preoperative BMI and weight loss after bariatric surgery in a total of 24,326 patients. It was noted that there appeared to be a negative association (less postoperative weight loss in patients with higher preoperative BMI) especially for studies using %EWL. The opposite was true for those studies reporting weight loss using absolute BMI instead. It seemed as if the outcome measure used was more predictive for the postoperative weight loss than the actual preoperative BMI [9].

There are other reasons as well to be aware of this variation. Lighter patients will have better %EWL and %EBMIL results, just because of this arithmetic effect. This mathematical bonus could well make the difference if applied to the 50 %EWL criterion for success for example. Surgeons can profit from this bonus as well. Operating mainly on the safer, lower BMIs will bring them better %EWL and %EBMIL results automatically.

The disappearance of variation by initial BMI does not only mean that it is caused only by the relative outcome measure used but also that it is not caused by gastric bypass

surgery. In other words, initial BMI does not influence the relative nadir weight loss result after gastric bypass: the operation works as effectively for heavier patients as for lighter patients.

The variation by initial BMI disappears when results of the BOLD data are expressed in relative measures with reference body mass 9 or 10 kg/m<sup>2</sup>. This gives a nice insight in the underlying mechanism of the bariatric procedure, as it means that on an average gastric bypass potentially reduces the part of the body mass above 10 kg/m<sup>2</sup> in all female patients by 49.6 %, regardless of their initial BMI. What could this mean? Undoubtedly, part of a person's body mass is not subjected to weight loss. This is the inert body mass of bones, tendons, nervous system, blood vessels, etc. Extrapolating loss hypothetically to the extreme, it becomes evident that losing weight is not held at ideal weight (%EWL) or at BMI 25 kg/m<sup>2</sup> (%EBMIL), nor that it can continue to zero (%TWL). From studies on extreme starvation in anorexia nervosa [10] and victims of famine [11], it is known that weight loss in fact stops at a minimal viable BMI of about 10 kg/m<sup>2</sup>, which matches the optimal reference point found in this study. However, this new relative measure probably is of no practical use for reporting bariatric outcome. Different people will have a different inert body mass depending on gender, race, and "body frame" and therefore might have different optimal reference points as well. This underlines that using relative measures in bariatric surgery is a bit like playing with fire. Probably %TWL, not having a reference point at all, has the least disadvantages with the least chance of burning your fingers. Of course, %EWL and %EBMIL do have a kind of emotional advantage above %TWL in that they imply a personal goal or target to be reached (a patient's ideal weight or normal BMI). On an individual basis, %EWL and %EBMIL results might therefore be more meaningful for some patients and their surgeons than %TWL results. This does not however outweigh the presented disadvantages that arise when these outcome measures are being used to compare weight loss not to a preset target but to the results of other patients for scientific purposes. For comparing patients with other patients in a scientific way, %TWL definitely is the better alternative, for several reasons. The BOLD data show that %TWL generates the smallest variation of the three. The differences in medians between subgroups A and B for %EBMIL and %EWL are 18.6 and 11.8 % respectively, which are in sharp contrast with the 2.0 % for %TWL. The variation by initial BMI is therefore more likely to interfere with the significance of conclusions on bariatric outcome using %EWL and %EBMIL than using %TWL. Moreover, unlike %EWL and %EBMIL, the variation generated by %TWL is not inverted to the variation in absolute weight loss: Heavier patients show more absolute weight loss and higher %TWL results (but lower %EWL and %EBMIL results). %TWL is used in

many bariatric reports in recent years; it is easy to comprehend and visualize and easy to calculate as well, because, as a patient's body length does not change during the weight loss process, percentage total weight loss will always equal percentage total BMI loss. Furthermore, it can be calculated from kilograms, pounds or BMI points alike, without the need for conversion. Finally, %TWL is compatible with what our colleagues in metabolic medicine use. For good reasons, they would not use %EWL or %EBMIL, as there is no connection what so ever with these outcome measures and existing criteria for obesity related health risk and metabolic impairment.

In summary, there is now strong evidence for abandoning %EWL and %EBMIL altogether. It should be recommended to express bariatric relative outcome only in %TWL, together with absolute kilograms or BMI points lost in all cases.

**Conflict of interest** The author Arnold van de Laar declares that he has no conflict of interest.

## References

- Dixon JB, McPhail T, O'Brien PE. Minimal reporting requirements for weight loss: current methods not ideal. *Obes Surg*. 2005;15(7):1034–9.
- Reinhold R. Critical analysis of long term weight loss following gastric bypass. *Surg Gynecol Obstet*. 1981;155:385–94.
- Higa KD, Boone KB, Ho T, et al. Laparoscopic Roux-en-Y gastric bypass for morbid obesity. Technique and preliminary results of our first 400 patients. *Arch Surg*. 2000;135:1029–34.
- Biertho L, Steffen R, Ricklin T, et al. Laparoscopic gastric bypass versus laparoscopic adjustable gastric banding: a comparative study of 1,200 cases. *J Am Coll Surg*. 2003;197(4):536–44.
- Biron S, Hould FS, Lebel S, et al. Twenty years of biliopancreatic diversion: what is the goal of the surgery? *Obes Surg*. 2004;14(2):160–4.
- van de Laar A, de Caluwé L, Dillemans B. Relative outcome measures for bariatric surgery. Evidence against excess weight loss and excess body mass index loss from a series of laparoscopic Roux-en-Y gastric bypass patients. *Obes Surg*. 2011;21(6):763–7.
- 1983 Metropolitan Height and Weight Tables. New York: Metropolitan Life Foundation. *Statistical Bulletin* 1983;64(1):2–9.
- The Committee on Standards for Reporting Results, Mason EE, Amaral J, Cowan GS, Deitel M, Gleysteen JJ, Oria HE. Standards for reporting results. *Obes Surg*. 1994;4(1):56–65.
- Livhits M, Mercado C, Yermilov I, et al. Preoperative predictors of weight loss following bariatric surgery: systematic review. *Obes Surg*. 2012;22(1):70–89.
- Rosling AM, Sparén P, Norring C, et 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.
- Collins S. The limit of human adaption to starvation. *Nat Med*. 1995;1(8):810–4.