OBES SURG (2011) 21:763–767
DOI 10.1007/s11695-010-0347-0

CLINICAL RESEARCH

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

Arnold van de Laar · Laura de Caluwé · Bruno Dillemans

Published online: 1 January 2011
© Springer Science+Business Media, LLC 2010

Abstract
Background Bariatric results expressed in the relative measure excess weight loss (%EWL) vary significantly by initial body mass index (BMI): the heavier the patient, the lower the %EWL. We examine if this variation is caused by using a wrong outcome measure and argue that no relative weight loss measure can express bariatric or metabolic goals unequivocally.
Methods Nadir weight loss results after laparoscopic gastric bypass in 168 women with initial BMI ≥35 to <60 kg/m^2 are calculated for %EWL and 61 different relative measures using the formula 100%×(initial BMI−nadir BMI)/(initial BMI−a), with a ranging from −30 to +30. Standard deviations are compared mutually and with those reported in the literature. For each relative measure, the significance of any variation by initial BMI is determined with the Mann–Whitney U test.
Results Mean initial BMI was 44.9±6.7 (35.0–59.7)kg/m^2. Mean nadir BMI was 28.8±5.8 (18.5–44.4)kg/m^2. Mean nadir excess BMI loss (%EBL; a=25) was 87.0±28.0 (19.4–155.1)%. Mean nadir (total) weight loss (%TWL; a=0) was 35.9±8.5 (9.5–57.1)%. Mean nadir %EWL was 77.3±22.8 (17.7–135.2)%. The smallest variation coefficient was 23.7% at a ranging from −1 to +3, including %TWL (a=0). This is lower than variation coefficients of %EWL results in our series and in the literature. Variation by initial BMI is significant using relative measures with a≥3, including %EBL and %EWL (both p<0.0001) and not significant with a<3, including %TWL (p=0.13).
Conclusions In contrast to their widespread use, %EBL and %EWL are not suited for comparing different patients or nonrandomized groups. They cause variation by initial BMI, which disappears using %TWL. In general, absolute terms should be preferred for bariatric outcome and goals. The power of bariatric procedures is best represented by their mean %TWL value.

Keywords Bariatric surgery · Outcome measure · Weight loss measure · Excess weight loss · Excess BMI loss · Gastric bypass · Bariatric goal · 50%EWL

Introduction

Weight loss results are expressed in absolute terms like kilogram or body mass index (BMI) and relative terms like percentage excess weight loss (%EWL), percentage excess BMI loss (%EBL), or percentage (total) weight loss (%TWL). A recent multivariable study showed %EWL to be substantially less accurate as an outcome measure for bariatric surgery than absolute weight [1]. Absolute terms are favored by physicians in nonsurgical reports on weight loss. Relative measures are used mainly by surgeons [2]. They do this for three reasons.

1. 50%EWL is used as a mark between failure and success for bariatric surgery. It originates from misquoting the 1982 Reinhold criteria [3]. But although

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

L. de Caluwé · B. Dillemans
Algemeen Ziekenhuis Sint-Jan,
Ruddershove 10,
8000 Brugge, Belgium

not evidence-based, 50%EWL can be useful as an arbitrary milestone on the way to a bariatric goal, as long as it is not mistaken for the goal of bariatric surgery itself. Like for all treatments of any disease, the goal of bariatric surgery is twofold: health risk reduction and improvement of health-related quality of live (QoL). Metabolic surgery adds a third. These three goals might be achieved at 50%EWL, but not necessarily. First, there is a direct correlation between health risk and BMI [4]. As %EWL depends on a patient's initial BMI, One single % EWL value can correspond with a wide variety of possible BMI results in different patients. Therefore, %EWL is not able to express health risk reduction unequivocally in different patients. Second, many factors other than weight loss play an important role in postoperative QoL [5]. Therefore, %EWL is not able to express improvement of QoL unequivocally. Third, obesity in the metabolic syndrome is measured not by weight but by waist circumference [6]. Therefore, %EWL is not able to express metabolic improvement. The reasoning in these three deductions concerns all relative weight loss measures, not only %EWL.

2. Relative measures are used to express individual weight loss. Relative measures are dependent of a patient's initial BMI. Using them implies that the effect of bariatric surgery is in some way dependent of a patient's initial BMI as well. If both dependencies match, bariatric results would show an even distribution in a group of patients with different initial BMI. There is evidence that this is not the case for %EWL. In large series reporting weight loss outcome of bariatric surgery in the past decade, results expressed in %EWL were found to vary significantly if grouped by initial BMI: The heavier the patient, the smaller the %EWL (and vice versa, the lighter the patient, the higher the % EWL). This was observed first for laparoscopic gastric bypass in 2000, for gastric banding in 2003, and for biliopancreatic diversion in 2004 [7–9]. Although these findings were clear, they did not draw much attention, and whether this initial BMI-driven variation is inherent to bariatric surgery or is caused by using a wrong measure has not yet been examined.
3. The mean relative weight loss result in a series of patients undergoing a specific bariatric procedure, by a specific surgeon, or in a specific center, is used to indicate the power of that procedure, or that surgeon, or that center. A higher mean relative result suggests a more powerful procedure, a better surgeon, or a better center. However, surgeons using mean %EWL can improve "their power" paradoxically by operating mainly on the safer, lower end of the BMI scale, as lighter patients show higher %EWL values.

To find a relative measure best suited for bariatric weight loss results and bariatric power, we look at variations by initial BMI and at the size of deviations in one set of bariatric results expressed in different relative measures.

Materials and Methods

Bariatric weight loss results can be expressed using different relative measures. Relative weight loss measures express weight loss relative to, first of all, the initial state and, in general, to a specified reference point (a) as well. In the case of %EWL, this reference point is the ideal weight, and in the case of %EBL, this point is BMI 25. Many hypothetical relative weight loss measures can be constructed using different reference points. Using some relative measures might lead to more deviation in results than using others. If so, some measures would be less suited for expressing bariatric outcome than others. The size of deviation could then be used to find the measure best suited for expressing weight loss results. For this purpose, we look at weight loss results of a group of patients with a wide range of patient characteristics, but with similar bariatric procedure characteristics. From a recently published single surgeon, single-center retrospective series of 2,606 patients undergoing a highly standardized laparoscopic fully stapled Roux-en-Y gastric bypass [10], a sample of 168 patients with minimum follow-up of 2 years is selected randomly with specified patient characteristics. All 168 subjects are Caucasian women. Subjects are distributed by initial BMI in four groups ($\geq$35 to <40 kg/m^2, $\geq$40 to <45 kg/m^2, $\geq$45 to <50 kg/m^2, and $\geq$50 to <60 kg/m^2) and in each BMI group by age in three subgroups (18–31, 32–45, and 46–59 years), making 12 subgroups of 14 patients, in order to aspire an even distribution of patient characteristics within the sample. Initial BMI and nadir BMI, defined as lowest BMI within 2 years postoperatively, are compared. These absolute data are then transformed into relative data. We calculate different relative nadir weight loss results with the formula derived from %EBL: $100\% \times (\text{initial BMI} - \text{nadir BMI})/(\text{initial BMI} - a)$, in which the reference point a is any whole number ranging from −30 to +30, resulting in 61 datasets of 168 results each. An extra set is calculated for %EWL using mean values for medium body frame from the Metropolitan Life Insurance Company Tables for Ideal Weight [11]. Whether one of these 62 relative measures is more or less meaningful for use in bariatric surgery than the others depends on the accuracy of the bariatric outcome expressed in this relative measure, compared to using the other measures. We determine this accuracy in two ways: by comparing the deviations in the results and by measuring the variation by initial BMI within the results. Therefore, we perform with these 62 datasets two oper-

ations. First, for each of these 62 datasets, mean and standard deviation (SD) are determined. As it can be expected that these datasets of ratio variables have widely different means, we compare variation coefficients (VC= 100%×SD/mean) instead of SD in order to interpret the respective SD in the context of the mean. Different calculations of the same set of bariatric outcome results are compared, not different bariatric outcome results mutually. Therefore, any difference in VC can be considered significant. Furthermore, we searched PubMed for studies published since 2000 reporting SD on %EWL results in laparoscopic gastric bypass series with $n \geq 100$ patients at 1- or 2-year follow-up and compare them with our results for historical significance.

Second, to demonstrate variation by initial BMI, the sample of 168 subjects is divided into two groups of 84 by initial BMI: A, ≥35 to <45 kg/m^2; and B, ≥45 to <60 kg/m^2. Statistic significance of any difference in results between both groups is determined with the Mann–Whitney U test, considering a two-tailed $p<0.05$ significant.

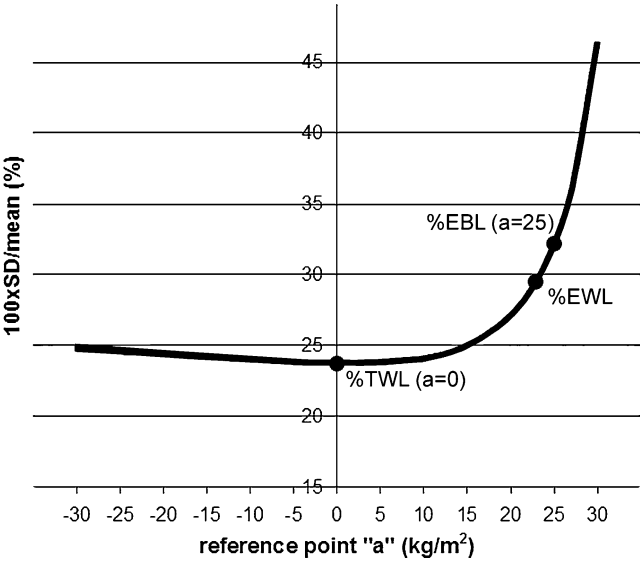

Fig. 2 Variation coefficient (100% × SD/mean) for results expressed in different relative measures with formula 100%×(initial BMI−nadir BMI)/(initial BMI−a) with reference point a ranging from −30 to +30, and in percentage excess weight loss (%EWL)

Results

Initial BMI and nadir BMI are shown in Fig. 1. Results are reported as median/mean±1SD (range). Age was 36/37.6± 10.4 (18–59)years. Initial BMI was 44.9/44.9±6.7 (35.0–59.7)kg/m^2. Nadir BMI was 28.4/28.8±5.8 (18.5–44.4)kg/m^2. Nadir %EBL was 85.5/87.0±28.0 (19.4–155.1)%. Nadir % TWL was 36.0/35.9±8.5 (9.5–57.1)%. Nadir %EWL was 77.2/77.3±22.8 (17.7–135.2)%. Figure 2 shows the VC for results expressed in relative measures with reference point a ranging from −30 to +30 and %EWL. The smallest VC is 23.7% and is found at a ranging from −1 to +3. The VC for %EBL (a=25) is 32.2%, for %EWL 29.4%, and for %TWL (a=0) 23.7%. A PubMed search yielded seven studies published since 2000 reporting mean %EWL with SD on $n \geq 100$ patients at 1 to 2 years after laparoscopic gastric bypass [12–18] (Table 1). The mean VC of the % EWL results in these reports is 29%, which equals the VC of %EWL results in our series. The difference in results between the lower and the higher initial BMI groups A and B ("variation by initial BMI") is significant ($p<0.05$) using %EWL and relative measures with $a \geq 3$ (including %EBL). It is not significant ($p>0.05$) using relative measures with $a<3$ (including %TWL). For %EWL, $p<0.0001$ (U_A=1,244, z=7.20). For a=25 (%EBL), $p<0.0001$ (U_A=961, z=8.14). For a=0 (%TWL), p=0.13 (U_A=3,049, z=1.52).

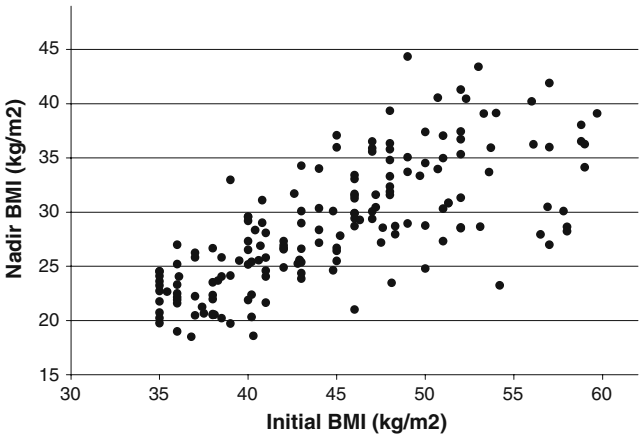

Fig. 1 Initial BMI and nadir BMI (defined as lowest BMI within 2 years postoperatively) in a sample of 168 Caucasian women who underwent a standardized, single surgeon, fully stapled laparoscopic Roux-en-Y gastric bypass

Discussion

1. The highly significant ($p<0.0001$) variation by initial BMI of weight loss results expressed in %EBL and %EWL in our study confirms findings that outcome of bariatric surgery expressed in %EWL varies if grouped by initial BMI. In our study, this variation disappears when expressing the same results in a different relative measure. This is to our knowledge the first evidence that this variation is caused by using the wrong bariatric measure and that it is not intrinsic to any bariatric mechanism. It should therefore be concluded that, although %EBL and %EWL can be used for individual weight loss results, they are not suited for comparing patients with different initial BMI or nonrandomized groups.

Table 1 Studies published since 2000 reporting mean %EWL including SD on $n \geq 100$ patients at 1 to 2 years after laparoscopic gastric bypass

Authors	*n* (at 1–2-year follow-up)	Year of publication	Mean %EWL	Standard deviation	Variation coefficient
Schauer et al. [12]	101	2000	68.8	22	32
Sekhar et al. [13]	477	2007	71.3	18.4	25.8
Jan et al. [14]	237	2007	64.9	18	28
Strain et al. [15]	101	2009	70.4	19.7	28.0
Christou et al. [16]	315	2009	70.4	22.5	32.0
Welch et al. [17]	100	2010	59.1	17.2	29.1
Kim et al. [18]	219	2010	66.2	19.7	29.8
Total 7	1,550	Mean (7)	67.3	20	29
		Mean (1,550)	68.4	20	29

2. If using one measure causes more deviation than necessary, the size of deviation can be used to find a better measure. In our study, results expressed in % TWL show the lowest VC of all relative measures tested and without any significant variation by initial BMI (p=0.13). This makes %TWL the better relative measure for reporting bariatric outcome, suited for comparing both different patients and different series and reports. Several recommendations for reporting weight loss in bariatric surgery have been published since 2003 leading away from the use of % EWL [19–21]. Main criticism was the ambiguity of %EWL using a reference point that differs for each patient. %EBL was proposed as better alternative and scientifically more correct as it is based on one single reference point. However, it does not solve the problem of variation by initial BMI. This was noticed by Baltasar et al. in 2008 who suggested to make some corrections using a sliding scale of reference points for %EBL by initial BMI [22]. Again, however, this would result in a measure with different reference points for each patient. We show that variation by initial BMI can be avoided all together by using %TWL, resulting in one relative measure suited for all patients. This is not only a practical consequence of our findings. It touches the base of what bariatric surgeons are doing: obviously, the weight loss mechanism after bariatric surgery does not involve a reference point.
3. The main advantage of a new outcome measure must be simplification of the existing international measures. %TWL does not only generate the narrowest deviation without causing variation by initial BMI, it has some specific advantages as well, making it indeed simpler to use than other relative measures. %TWL is easier to calculate, comprehend, and explain to the patient than %EBL or %EWL, and the patient can easily translate the %TWL value into actual weight, as put forward by Karmali et al. in 2009 [23]. Also, with increasing insight in metabolic effects of bariatric surgery, it is important to improve communication with endocrinologists, who mainly deal with outcome in absolute terms. % TWL relates better to absolute terms than other relative measures because no other reference point has to be considered than the initial state of the patient. Besides, % EWL is not part of the endocrinologist's vocabulary. Furthermore, as a person's body length does not change during the weight loss process, %TWL will always equal percentage (total) BMI loss (%TBL).
4. %EWL "behaves" like a relative measure with reference a=23, with comparable, significant variation by initial BMI and VC. %EBL (a=25) causes a greater VC and is therefore not a better alternative for %EWL.
5. No relative weight loss measure is suited to express or define any of the goals of bariatric and metabolic surgery. They might be useful to express, by consensus, an arbitrary level of weight loss success, but without claiming a goal. Absolute terms should be preferred for bariatric and metabolic goals and for individual weight loss results in conformity with common practice in nonsurgical reports on weight loss.
6. A narrower SD is observed using %TWL than found in reports using %EWL. Bariatric surgery might therefore be more predictable than it seems. Mean %TWL of a series of patients who underwent a specific bariatric procedure is better suited than mean %EWL to represent the power of that procedure. Mean nadir % TWL in our series was 35.9±8.5%. This might be used to advise patients in choosing a tailor-made procedure. For example, envisioning a health risk reduction for a specific female patient aiming at a nadir BMI below 35 kg/m^2, we would expect her chance to reach this goal with our laparoscopic gastric bypass to be less than one in two if her initial BMI is above 55 kg/m^2, or less than one in six if it is above 60 kg/m^2. We might therefore suggest more restriction or malabsorption from the beginning and discuss with her the options of a (adjustable) banded gastric bypass or a biliopancreatic diversion as primary procedure.

7. Our conclusion on preferring %TWL above %EBL and % EWL is based on mathematical evidence. Our mathematical method can be applied by any surgeon on his or her own results to come to similar conclusions. This might convince to change the way we report our outcome.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Dallal RM, Quebbemann BB, Hunt LH, et al. Analysis of weight loss after bariatric surgery using mixed-effects linear modeling. Obes Surg. 2009;19:732–7.
2. Dixon J, McPhail T, O'Brien P. Minimal reporting requirements for weight loss: current methods not ideal. Obes Surg. 2005;15 (7):1034–9.
3. Reinhold R. Critical analysis of long-term weight loss following gastric bypass. Surg Gynecol Obstet. 1982;155:385–94.
4. Whitlock G, Lewington S, Sherliker P, et al. Body-mass index and cause-specific mortality in 900, 000 adults: collaborative analyses of 57 prospective studies. Lancet. 2009;373(9669):1083–96.
5. van Hout G, van Heck G. Bariatric psychology, psychological aspects of weight loss surgery. Obes Facts. 2009;2(1):10–5.
6. Alberti K, Zimmet P, Shaw J. IDF Epidemiology Task Force Consensus Group. The metabolic syndrome—a new worldwide definition. Lancet. 2005;366(9491):1059–62.
7. 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.
8. 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.
9. Biron S, Hould F, Lebel S, et al. Twenty years of biliopancreatic diversion: what is the goal of the surgery? Obes Surg. 2004;14:160–4.
10. Dillemans B, Sakran N, Van Cauwenberge S, et al. Standardization of the fully stapled laparoscopic Roux-en-Y gastric bypass for obesity reduces early immediate postoperative morbidity and mortality: a single centre study on 2606 patients. Obes Surg. 2009;19(10):1355–64.
11. Metropolitan Height and Weight Tables. New York: Metropolitan Life Foundation. Statistical Bulletin. 1983;64(1):2–9.
12. Schauer PR, Ikramuddin S, Gourash W, et al. Outcomes after laparoscopic Roux-en-Y gastric bypass for morbid obesity. Ann Surg. 2000;232(4):515–29.
13. Sekhar N, Torquati A, Youssef Y, et al. A comparison of 399 open and 568 laparoscopic gastric bypasses performed during a 4-year period. Surg Endosc. 2007;21:665–8.
14. Jan JC, Hong D, Bardaro SJ, et al. Comparative study between laparoscopic adjustable gastric banding and laparoscopic gastric bypass: single-institution, 5-year experience in bariatric surgery. Surg Obes Relat Dis. 2007;3:42–51.
15. Strain GW, Gagner M, Pomp A, et al. Comparison of weight loss and body composition changes with four surgical procedures. Surg Obes Relat Dis. 2009;5:582–7.
16. Christou N, Efthimiou E. Five-year outcomes of laparoscopic adjustable gastric banding and laparoscopic Roux-en-Y gastric bypass in a comprehensive bariatric surgery program in Canada. Can J Surg. 2009;52(6):249–58.
17. Welch G, Wesolowski C, Zagarins S, et al. Evaluation of clinical outcomes for gastric bypass surgery: results from a comprehensive follow-up study. Obes Surg. 2010. doi:10.1007/s11695-009-0069-3.
18. Kim S, Richards WO. Long-term follow-up of the metabolic profiles in obese patients with type 2 diabetes mellitus after Roux-en-Y gastric bypass. Ann Surg. 2010;251(6):1049–55.
19. Deitel M, Greenstein RJ. Recommendations for reporting weight loss (Editorial). Obes Surg. 2003;13:159–60.
20. Deitel M, Gawdat K, Melissas J. Reporting weight loss 2007 (Editorial). Obes Surg. 2007;17:565–8.
21. Greenstein R. Reporting weight loss (Letter). Obes Surg. 2007;17:1275–6.
22. Baltasar A, Deitel M, Greenstein RJ. Weight loss reporting. Obes Surg. 2008;18(6):761–2. Epub 2008 Apr 12.
23. Karmali S, Birch DW, Sharma AM. Is it time to abandon excess weight loss in reporting surgical weight loss? Surg Obes Relat Dis. 2009;4:503–6.