A Report on "Once-Weekly Semaglutide in Adults with Overweight or Obesity" by Wilding et al. (2021)

Reviewer 2

February 06, 2026

v1

isitcredible.com

Disclaimer

This report was generated by large language models, overseen by a human editor. It represents the honest opinion of The Catalogue of Errors Ltd, but its accuracy should be verified by a qualified expert. Comments can be made here. Any errors in the report will be corrected in future revisions.

I am wiser than this person; for it is likely that neither of us knows anything fine and good, but he thinks he knows something when he does not know it, whereas I, just as I do not know, do not think I know, either. I seem, then, to be wiser than him in this small way, at least: that what I do not know, I do not think I know, either.

Plato, *The Apology of Socrates*, 21d

To err is human. All human knowledge is fallible and therefore uncertain. It follows that we must distinguish sharply between truth and certainty. That to err is human means not only that we must constantly struggle against error, but also that, even when we have taken the greatest care, we cannot be completely certain that we have not made a mistake.

Karl Popper, 'Knowledge and the Shaping of Reality'

Overview

Citation: Wilding, J.P.H., Batterham, R.L., Calanna, S., Davies, M., Van Gaal, L.F., Lingvay, I., McGowan, B.M., Rosenstock, J., Tran, M.T.D., Wadden, T.A., Wharton, S., Yokote, K., Zeuthen, N., and Kushner, R.F. (2021). Once-Weekly Semaglutide in Adults with Overweight or Obesity. *New England Journal of Medicine*. Vol. 384, No. 11, pp. 989–1002.

Abstract Summary: A double-blind trial of 1961 adults with overweight or obesity without diabetes found that once-weekly subcutaneous semaglutide (2.4 mg) plus lifestyle intervention resulted in a mean body weight reduction of 14.9% at 68 weeks, compared to 2.4% with placebo plus lifestyle intervention.

Key Methodology: Randomized, double-blind, placebo-controlled trial (STEP 1) over 68 weeks, comparing once-weekly subcutaneous semaglutide (2.4 mg) to placebo, both with lifestyle intervention.

Research Question: Can adults with obesity achieve weight loss with once-weekly semaglutide at a dose of 2.4 mg as an adjunct to lifestyle intervention?

Summary

Is It Credible?

This article reports the results of the STEP 1 trial, a large, double-blind, randomized, placebo-controlled phase 3 study evaluating the efficacy and safety of once-weekly subcutaneous semaglutide (2.4 mg) for weight management. The authors conclude that among adults with overweight or obesity (without diabetes), the treatment was associated with a "sustained, clinically relevant reduction in body weight" (p. 989). Specifically, the study reports a mean change in body weight of -14.9% in the semaglutide group compared with -2.4% in the placebo group at week 68 (p. 989). Beyond weight loss, the article claims improvements in cardiometabolic risk factors and physical functioning, noting that participants receiving the drug had "a greater increase in participant-reported physical functioning" (p. 989).

The primary claim regarding the magnitude of weight loss appears highly credible. The trial design is robust, utilizing a large sample size of 1,961 participants and standard statistical methods to handle missing data (p. 990). The effect size—a difference of 12.4 percentage points between groups—is substantial enough that it is unlikely to be an artifact of minor methodological flaws or statistical noise. The objective nature of the primary endpoint (weight measured in kilograms) further insulates this central finding from bias. However, the interpretation of this weight loss as "sustained" requires careful qualification. The data presented covers only the on-treatment period up to week 68. Although the study protocol included a 7-week off-treatment follow-up period, and the participant flow chart indicates that nearly 95% of participants attended the final week 75 visit, the article omits efficacy data from this period (p. 28 of Supplementary Appendix). Without evidence of what occurred after the drug was withdrawn, the claim of "sustained" reduction is valid only within the context of continued therapy; it does not establish durability post-cessation.

The credibility of the secondary claims regarding participant-reported outcomes (PROs), such as improvements in physical functioning scores (SF-36 and IWQOL-Lite-CT), is weaker due to the potential for functional unblinding. While the study was double-blind, the side effect profile was distinct: 74.2% of participants in the semaglutide group reported gastrointestinal disorders compared to 47.9% in the placebo group (p. 997). Nausea alone was reported by 44.2% of the treatment group versus 17.4% of the placebo group (p. 998). It is plausible that many participants correctly guessed their assignment based on these adverse events. Unlike body weight, PROs are subjective; knowledge of treatment assignment, combined with the psychological boost of visible weight loss, could inflate reported improvements in quality of life.

Furthermore, the generalizability of the findings is constrained by the study's specific inclusion and exclusion criteria. The cohort was predominantly female (74.1%) and White (75.1%), and the trial excluded individuals with type 2 diabetes, significant mental health history, or recent weight fluctuations (pp. 990, 992; p. 6 of Supplementary Appendix). While these exclusions are standard for isolating efficacy and ensuring safety in a clinical trial, they result in a "clean" population that may not fully represent the complex, multi-morbid patient profiles often encountered in clinical practice. Finally, the extensive involvement of the sponsor, Novo Nordisk—who designed the trial, oversaw conduct, and analyzed the data—warrants a degree of caution regarding the framing of the results, particularly the optimistic interpretation of body composition changes based on a small, non-representative sub-study of only 140 participants (pp. 990, 1000). Despite these caveats, the core finding of significant weight loss during treatment remains solid.

The Bottom Line

The claim that once-weekly semaglutide 2.4 mg leads to substantial weight loss (~15%) in adults with overweight or obesity is highly credible and supported by robust trial data. However, the assertion that this weight loss is "sustained" is only proven while the patient remains on the drug; the article does not present data on weight trajectory after treatment cessation. Additionally, improvements in subjective quality-of-life metrics should be viewed with caution due to the likelihood that side effects revealed treatment assignment to participants.

Potential Issues

Omission of efficacy data from the off-treatment follow-up period: The study protocol included a 7-week off-treatment follow-up period from week 68 to week 75, and the participant flow chart confirms that 94.9% of the semaglutide group and 93.0% of the placebo group attended the final week 75 visit (p. 28 of Supplementary Appendix). However, the article presents no efficacy data from this period. The primary and confirmatory endpoints were pre-specified at week 68, and the follow-up period was primarily for safety monitoring (p. 991). Nonetheless, the absence of the collected week 75 efficacy data is a significant omission. It prevents an assessment of the durability of weight loss after treatment cessation, a critical question for patients and clinicians regarding the potential for weight rebound. This omission limits the interpretation of the article's conclusion of a "sustained, clinically relevant reduction in body weight," as the term "sustained" cannot be fully evaluated without data on what occurs after the drug is withdrawn (p. 989).

Limited generalizability due to a highly selected patient population: The study's findings may have limited applicability to the broader, more complex population of individuals with overweight or obesity seen in typical clinical practice. The authors acknowledge several of these limitations in the discussion, including the "preponderance of women and White participants" and the "exclusion of persons with type 2 diabetes" (p. 1000). The overall study population was 74.1% female and 75.1% White (p. 993). Furthermore, the exclusion criteria were extensive, removing patients with diabetes, a history of chronic pancreatitis, and prior surgical obesity treatment (p. 990). The trial also excluded individuals with significant mental health conditions, such as a recent history of major depressive disorder or a lifetime history of a suicidal attempt (p. 6 of Supplementary Appendix). While these exclusions are common in clinical trials for safety and to isolate treatment effects, they result in a study conducted in a more homogeneous and potentially healthier cohort than the

multi-morbid population that would typically be considered for this medication.

Potential for bias due to sponsor funding and oversight: The study was designed, funded, and overseen by the drug's manufacturer, Novo Nordisk. The article is transparent about the sponsor's role, stating that the sponsor "designed the trial and oversaw its conduct... undertook site monitoring, data collation, and analysis" and that the manuscript was drafted with assistance from a "sponsor-funded medical writer" (p. 990). While this level of industry involvement is standard for large Phase 3 trials required for regulatory approval, and the academic authors vouch for the data's accuracy, this arrangement represents a structural conflict of interest. The sponsor's involvement in every critical stage of the research process creates a potential for bias in the framing and interpretation of results that readers should consider when evaluating the study's conclusions.

Potential for unblinding to bias subjective outcomes: The study was designed as double-blind, but the distinct and high rate of gastrointestinal side effects in the semaglutide group may have led to functional unblinding of participants and investigators. Gastrointestinal disorders were reported in 74.2% of participants receiving semaglutide compared to 47.9% receiving placebo (p. 997). While the primary endpoint of body weight is objective and robust to this potential bias, several secondary endpoints were subjective patient-reported outcomes (PROs), such as the SF-36 and IWQOL-Lite-CT scores. It is possible that participants who correctly guessed they were receiving the active drug due to its side effects may have reported more favorable subjective outcomes. The article does not discuss this potential source of bias or report any assessment of blinding effectiveness.

Non-representative sample for body composition sub-study: The article's conclusions about changes in body composition, such as the finding that semaglutide preferentially reduces fat mass, are based on a small sub-study that may not be representative of the overall trial population. The authors acknowledge that these assessments "were performed in only a subpopulation of participants" (p. 1000). This

sub-study included only 140 individuals, or 7.1% of the total trial population. A comparison of the sub-study demographics with the overall population reveals that the sub-study group was older (mean age 51 vs. 46), had a lower mean BMI (34.8 vs. 37.9), and had a different racial composition, with substantially fewer Asian participants (1.4% vs. 13.3%) (p. 992; p. 17 of Supplementary Appendix). These differences may limit the extent to which the body composition findings can be generalized to the main trial population.

Uncertainty in an exploratory finding due to subjective baseline assessment: An exploratory finding that 84.1% of participants with prediabetes "reverted to normoglycemia" may be affected by the method used for the baseline diagnosis (p. 995). The article notes that "The presence of prediabetes was determined by investigators on the basis of available information... and in accordance with American Diabetes Association criteria" (p. 993). This method was not standardized with a uniform protocol across the 129 trial sites and relied on investigator judgment and the variable quality of available medical records. This introduces a risk of misclassification bias at baseline, which could affect the reliability of the reported reversion rate.

Minor presentation and documentation issues: There are minor issues in the presentation of baseline characteristics. First, the "Other" race category in Table 1 is difficult to interpret because it combines multiple distinct racial and ethnic groups with a methodological artifact, "'not applicable,' the last of which is the way race or ethnic group was recorded in France" (p. 993). The authors are transparent about this limitation, which stems from a regulatory constraint. Second, the footnote defining the "coexisting conditions" variable lists several conditions (e.g., cerebrovascular disease, gout) for which the baseline prevalence is not reported in the main table (p. 993). This represents a minor documentation gap in the characterization of the study population's baseline morbidity.

Future Research

Evaluation of post-treatment durability: Future research should prioritize the collection and reporting of efficacy data following the withdrawal of the medication. To validate the "sustained" nature of the weight loss, studies must track participants for a significant period (e.g., 6–12 months) after the cessation of therapy to quantify the rate and magnitude of weight regain.

Assessment of blinding efficacy: To address the risk of functional unblinding in trials of GLP-1 receptor agonists, future study designs should include a formal assessment of blinding success, asking participants to guess their treatment assignment at the trial's conclusion. This would allow researchers to statistically adjust subjective patient-reported outcomes (such as physical functioning scores) for potential unblinding bias.

Validation in diverse populations: Future trials should specifically target demographic groups underrepresented in this study, particularly men and non-White populations, as well as individuals with common comorbidities excluded here, such as controlled depression or history of bariatric interventions. This is necessary to determine if the efficacy and safety profile remains consistent across the broader patient population likely to be prescribed the drug.

isitcredible.com