Received: 12 April 2019 | Revised: 16 May 2019 | Accepted: 28 May 2019

DOI: 10.1111/dom.13800

ORIGINAL ARTICLE

WILEY

A pre-specified statistical analysis plan for the VERIFY study: Vildagliptin efficacy in combination with metformin for early treatment of T2DM

David R. Matthews MD[1,2] | Päivi M. Paldánius PhD[3] | Michael Stumvoll MD[4] | Jackie Han MS[5] | Giovanni Bader PhD[3] | YannTong Chiang PhD[5] | Pieter Proot PharmD[3] | Stefano Del Prato MD[6]

[1]Oxford Centre for Diabetes Endocrinology and Metabolism, Radcliffe Department of Medicine, Oxford, UK

[2]Harris Manchester College, University of Oxford, Oxford, UK

[3]Department of Cardiovascular Metabolism, Novartis Pharma AG, Basel, Switzerland

[4]Divisions of Endocrinology and Diabetes, University Hospital Leipzig, Leipzig, Germany

[5]Clinical Development and Analytics, Novartis Pharmaceutical Corporation, East Hanover, New Jersey

[6]Department of Clinical and Experimental Medicine, Section of Metabolic Diseases and Diabetes, University of Pisa, Pisa, Italy

Correspondence
David R. Matthews, MD, Harris Manchester College, Mansfield Road, Oxford, OX1 2TD, UK.
Email: david.matthews@ocdem.ox.ac.uk

Funding Information
Novartis Pharmaceutical Corporation

Peer Review
The peer review history for this article is available at https://publons.com/publon/10.1111/dom.13800.

Abstract

Aims: To ensure the integrity of the planned analyses and maximize the clinical utility of the VERIFY study results by describing the detailed concepts behind its statistical analysis plan (SAP) before completion of data collection and study database lock. The SAP will be adhered to for the final primary data analysis of the VERIFY trial.

Materials and Methods: Vildagliptin efficacy in combination with metformin for early treatment of T2DM (VERIFY) is an ongoing, multicentre, randomized controlled trial aiming to demonstrate the clinical benefits of glycaemic durability and glucose control achieved with an early combination therapy in newly-diagnosed type 2 diabetes (T2DM) patients.

Results: The SAP was initially designed at the study protocol conception phase and later modified, as reported here, in collaboration between the steering committee members, statisticians, and the VERIFY study leadership team. All authors were blinded to treatment allocation. An independent statistician has additionally retrieved and presented unblinded data to the independent data safety monitoring committee. An overview of the trial design with a focus on describing the fine-tuning of the analysis plan for the primary efficacy endpoint, risk of initial treatment failure, and secondary, exploratory and pre-specified subgroup analyses is provided here.

Conclusion: According to optimal trial practice, the details of the statistical analysis and data-handling plan prior to locking the database are reported here. The SAP accords with high-quality standards of internal validity to minimize analysis bias and will enhance the utility of the reported results for improved outcomes in the management of T2DM.

KEYWORDS

glycaemic durability, randomized controlled trial, statistical analysis plan, type 2 diabetes mellitus, VERIFY study, vildagliptin

1 | INTRODUCTION

The UK Prospective Diabetes Study was the first pivotal study that formed and defined a paradigm for expectations concerning therapeutic management of hyperglycaemia in individuals with newly diagnosed diabetes.[1] The study demonstrated positive microvascular outcomes with intensive glucose control policies as compared with conventional glucose control policies, and demonstrated improved outcomes in the subgroup of overweight patients receiving metformin, including decreased myocardial infarction, as well as improvements in diabetes-related endpoints and all-cause mortality.[2] However, the UK Prospective Diabetes Study relied on a stepwise approach based on metformin, sulfonylureas and insulin treatment. After completion of the study, new glucose-lowering agents that may offer more suitable early intensive treatment for individuals with type 2 diabetes mellitus (T2DM) have been introduced.

During the last decade, dipeptidyl peptidase-4 (DPP-4) inhibitors have been shown in various diverse populations,[3-6] in innovative randomized clinical trials[7] and also in non-interventional real-world studies,[8] to maintain improved glycaemic control without hypoglycaemia. Because of their main effect on the pancreatic islet cells, they have been recognized as ideal adjuncts to an insulin sensitizer such as metformin for an effective early treatment approach.[9] This allows the current treatment paradigm, which is characterized by ineffective lifestyle interventions, followed by monotherapy, which is often delayed, and the frequently prolonged periods of sustained hyperglycaemia that have become inevitable consequences of sequential clinical inertia, to be overcome.[10] Thus, it is appropriate to examine the role of a more intensive and effective initial therapy.

The VERIFY study was designed to investigate the long-term clinical benefits of early treatment intensification with a DPP-4 inhibitor (vildagliptin)-metformin combination over standard-of-care metformin monotherapy in maintaining durable glycaemic control in individuals with newly or recently diagnosed T2DM.[11] We hypothesized that a proactive approach of initiating early treatment with a pathophysiologically synergistic combination would increase the durability of glycaemic control as compared to a policy of prescribing metformin alone, followed by a second-line agent such as vildagliptin, only at the time of glycaemic deterioration.[12]

In contrast to many completed and ongoing cardiovascular (CV) outcome studies in those with established CV disease and/or multiple risk factors and long disease duration,[13] we aimed to recruit a representative diverse population reflecting the current characteristics of newly diagnosed persons at low CV risk. However, changes in the landscape of trials with an extensive focus on CV safety and some of the unexpected safety findings indicating increased risk of hospitalizations for heart failure (hHF) with saxagliptin in the SAVOR-TIMI trial,[14] led to an amendment to the VERIFY study protocol, stipulating now the inclusion of a CV adjudication committee and an unblinded independent data monitoring committee. Nevertheless, there have been no early indicators of unexpected or new safety findings during the study since 2012 or in any other populations with vildagliptin based on a recent holistic safety assessment exploring pharmacovigilance reports[15] or real-world evidence in diverse populations.[3,15]

2 | STUDY DESIGN

2.1 | Introducing the concept of early combination

The VERIFY study (Vildagliptin efficacy in combination with metformin for early treatment of T2DM; NCT01528254) is the first long-term study to address the concept of an early, initial combination treatment paradigm for T2DM. It is an ongoing, five-year, multinational, multi-ethnic study and the pragmatic study design and primary statistical assumptions have been described in detail in a previous publication.[11] The study comprises three treatment periods (Figure 1). During Period 1, patients were randomized to either early combination treatment with individualized doses of metformin (between 1 and 2 g/day) and vildagliptin (50 mg twice daily) or standard-of-care metformin monotherapy (+placebo). Randomization was performed using an interactive response technology (IRT) system. If initial treatment failed to maintain glycated haemoglobin (HbA1c) <7.0%, confirmed at two consecutive scheduled study visits, vildagliptin was added to metformin during Period 2. During Period 3, rescue therapy with insulin was acceptable for maintaining glycaemic control based on local guidelines, at the discretion of the investigator. The protocol also provided the option of choosing another third-line treatment strategy, leading to discontinuation from the study. The Period 3 start date is defined as the date of insulin initiation.

Both patients and investigators remained masked to treatment allocation during Period 1, and the study will compare the success of these two different approaches to treatment. All treatments are referred to as "treatment approach" throughout this statistical analysis plan (SAP) to acknowledge that patients switched from one treatment to another based on individualized response over the entire study period.

The main inclusion and exclusion criteria have been published previously.[11] Recruitment for the VERIFY trial commenced in March 2012 and, despite a narrow glycaemic range for inclusion, a linear rate of randomization of eligible, newly diagnosed patients worldwide was completed ahead of schedule in April 2014. The enrolment phase indicated that an absence of aptitude within healthcare systems, especially in developing countries, for identifying individuals at risk of diabetes was the major reason for screening failure, with approximately 80% of cases having an HbA1c value outside the protocol-defined, centrally assessed range. A total of 66 participants were classified as run-in failures because of metformin intolerance prior to uptitration to the lowest targeted dose of 1000 mg/day.

The last patient last visit was planned for early April 2019, followed by database lock in mid-May 2019. This manuscript, which describes the SAP, was submitted for publication in parallel with the last patient last visit and the planned database lock.

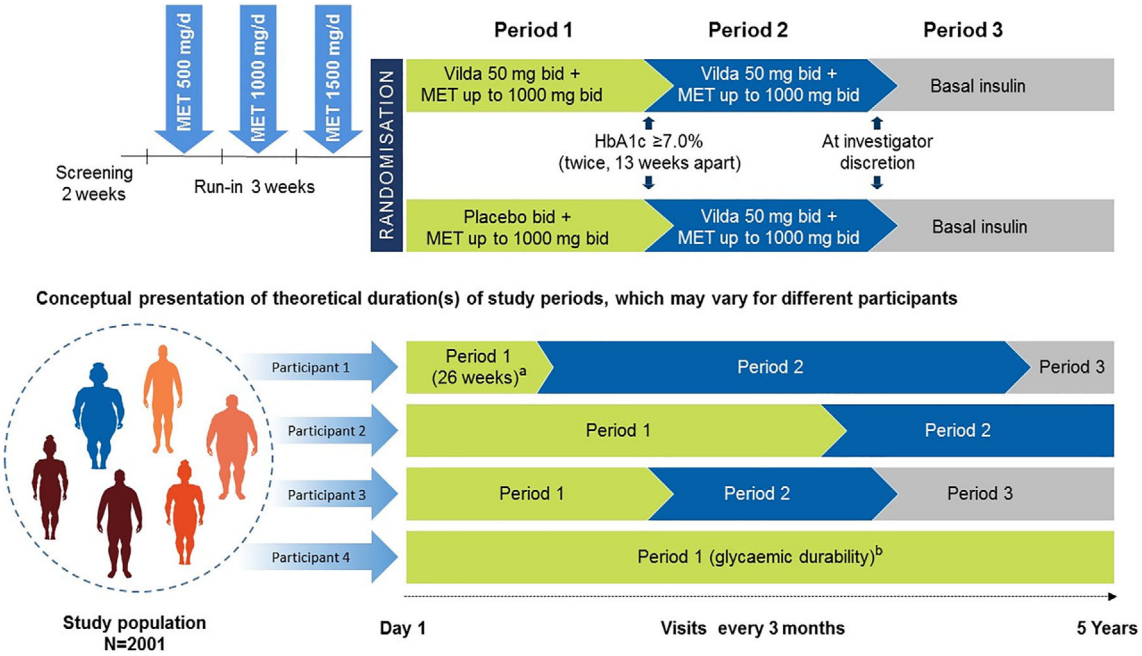

FIGURE 1 VERIFY study design. bid, twice daily; d, day; ECG, electrocardiogram; FPG, fasting plasma glucose; HbA1c, glycated haemoglobin; ISR/G, insulin secretion rate over glucose; MET, metformin; VILDA, vildagliptin. [a]Earliest progression from Period 1 to Period 2 can be at Week 26, when second of the two consecutive HbA1c levels is ≥7.0%, determined after ≥13 weeks of treatment. [b]Participants continuing to remain in Period 1 without the two consecutive HbA1c levels ≥7.0% for the entire study period demonstrate durability of glycaemic control. Similarly, not all participants will progress to Period 3 to receive insulin

2.2 | Baseline characteristics

Analysis of baseline characteristics was not included in the study protocol. However, as characterization of newly diagnosed T2DM patients has been considered in recent years to provide valuable information beyond the scientific study rationale, a post hoc analysis of pooled, blinded key characteristics was performed and has been published recently (Table 1).[16] Participants were generally young and well within the targeted glycaemic baseline range. Overall, the proportion of men and women was equal in the study. The presence of early microvascular complications was reported by 8% of the participants. Mean baseline estimated glomerular filtration rate (eGFR), assessed using the Modification of Diet in Renal Disease (MDRD) tool,[17] indicated normal renal function. The descriptive, blinded, country-level demographics and baseline characteristics were provided to all 34 participating countries and applicable regions for local publication.

3 | STATISTICAL ANALYSIS PLAN

3.1 | Primary endpoint and main analysis

The primary efficacy variable was time to confirmed initial treatment failure. Initial treatment failure is defined as HbA1c ≥7.0% at two consecutive scheduled visits, three months apart. Time to confirmed initial treatment failure is defined as the time from randomization until determination of the second of two consecutive HbA1c values ≥7.0% after at least 13 weeks of treatment, that is, at the end of Period 1, in line with the HbA1c target as defined by the main guidelines.[18]

Primary analysis concerning the primary efficacy variable will be performed in the full analysis set (FAS) using a Cox proportional hazard regression model with treatment approach and the pre-defined geographic region as classification variables and baseline HbA1c as a covariate. It will also be performed in the per-protocol set (PPS) as a supportive analysis (Table 2). Time to first treatment failure will be analysed similarly, to support the results of primary analysis. All analyses will be carried out at the one-sided 0.025 significance level (ie, two-sided 0.05).

3.2 | Sample size and statistical power

The planned sample size of 1000 per treatment approach will provide approximately 75% power to detect a risk reduction of 25% (ie, a hazard ratio of 0.75), with vildagliptin+metformin (vs metformin alone) as the primary variable. An annual initial treatment failure rate of 7.1% in the metformin monotherapy group is assumed,[19] along with an annual dropout rate of 4%, anticipated according to observations and data from this study. Statistical power for the primary variable was enhanced with two key modifications: (a) an expected lower dropout rate of approximately 4%, vs 11% as originally assumed,[19] and (b) one primary efficacy variable, vs two as in the original plan. Over the course of the study, smart and pragmatic inclusion of eligible participants with the ability to commit to the study for 5 years, based on an effective retention plan and continuing support from committed study site personnel, the final annualized discontinuation rate varied between only 4% and 5%. For the final approach, a single primary variable was adopted to optimize statistical power for the primary

TABLE 1 Key demographics and baseline characteristics of participants

Variable	Total
Patient population, n	2001
Gender: women, n (%)	1060 (53.0)
Age, years	
Median (IQR)	55 (48, 62)
Tertiles	51, 59
Race, n (%)	
White European	1217 (60.8)
Black	49 (2.4)
Asian	373 (18.6)
Native American	210 (10.5)
Other	152 (7.6)
Duration of T2DM, months	
Median (IQR)	3.4 (0.9, 10.3)
HbA1c, mmol/mol (%)	52 ± 3 (6.9 ± 0.3)
<7.0, n (%)	1423 (71.4)
≥7.0, n (%)	570 (28.6)
FPG, mmol/L	7.5 ± 1.5
Fasting insulin, median (IQR) (mU/L)	109 (75, 160)
HOMA-%β, median (IQR) (%)	84 (60, 116)
HOMA-%sensitivity, median (IQR) (%)	46 (31, 68)
BMI, kg/m^2	31.1 ± 4.7
<30 kg/m^2, n (%)	875 (43.7)
≥30 kg/m^2, n (%)	1125 (56.3)
GFR (MDRD), mL min^{-1} 1.73 m^{-2}	87.4 ± 18.5
Metformin daily dose, mg	1597.3 ± 396.5
1000 mg, n (%)	520 (26.3)
1500 mg, n (%)	678 (33.9)
2000 mg, n (%)	796 (39.8)

Note: ± indicates standard deviation.
Abbreviations: BMI, body mass index; BP, blood pressure; FPG, fasting plasma glucose; GFR, glomerular filtration rate; HbA1c, glycated haemoglobin; HOMA-β, homeostatic model assessment-%β; HOMA% sensitivity homeostatic model assessment % sensitivity; IQR, interquartile range; MDRD, modification of diet in renal disease; T2DM, type 2 diabetes mellitus.

analysis, with no alpha adjustment needed for the two primary endpoints as in the original approach.

3.3 | Handling of missing data

In the analysis for the primary efficacy variable, patients discontinuing the study for any reason during Period 1 were to be treated as censored values at the time of discontinuation. Patients who remained under the glycaemic threshold, or for whom a value above it was not confirmed at the next scheduled visit, were to be censored at the time of last study visit. No imputation will be made for missing HbA1c values when assessing endpoints. Visits for which HbA1c values are available will be used for assessment of the two consecutive visits for time to initial treatment failure. However, imputation will be used to avoid missing values for display of continuous variables, such as those for key secondary analysis parameters, as described below. For graphical presentation of continuous variables, such as HbA1c and FPG over time, the last observation carried forward (LOCF) approach has been used for systematic inclusion of all time-points for these variables for all participants, if applicable. Details of variable derivation are specified in the final SAP.

3.4 | Key secondary analyses

The rate of loss of glycaemic control over time will be estimated by an annualized slope of HbA1c values from Week 26 to the end of Period 1. Data from Week 26 to the end of Period 1 will be analysed using a linear mixed effect model, including treatment approach and geographic regions as classification variables, baseline HbA1c and time (in years) of HbA1c measurement as covariates and the interaction of treatment approach by time. An unstructured covariance will be used in the linear mixed effect model. The rationale for choosing the starting point for estimating the rate (slope) of loss of glycaemic control over time is based on previous observations,[12] in which the maximal reduction in mean HbA1c from baseline occurred between 24 and 32 weeks. Similar analyses will be performed to determine the rate of loss in glycaemic control from 26 weeks after study initiation until the end of Period 2.

Fasting plasma glucose (FPG) measurements at visits will be assessed similarly to determine the rate of loss of glycaemic control during the corresponding time periods. For those participants who consented to undergo a standardized meal-test, as well as C-peptide measurements at baseline, at 13 weeks and then annually, we will determine changes in the area under the curve (AUC) of insulin secretion rate over glucose (ISR/G), as an assessment of β-cell function, as well as changes in oral glucose insulin sensitivity (OGIS).

Analysis data sets for secondary endpoints will include either FAS, for variables for which all participants are observed over 5 years. For example, FPG development during Period 2 is applicable only to those who entered Period 2) and/or those who participated in the meal-test sub-study. All secondary analyses, including those pre-specified in this SAP, will be of an exploratory/descriptive nature andn thereforen no formal hypothesis testing will be performed on these, in order to avoid multiplicity. As planned, a statistical test for treatment comparison will be performed for each secondary endpoint using a pre-specified analysis model. Treatment difference, 95% confidence intervals and nominal *P* values are to be provided. The nominal *P* values, without further adjustment, that are associated with statistical assessment of the secondary analyses are provided as supportive evidence for scientific discussion and further hypothesis generation, but not as a definite claim. The confidence intervals for treatment estimates/differences will be useful in quantifying the expected treatment effects for comparison between the two treatment groups.

3.5 | Supportive descriptive analyses

Absolute values and change in HbA1c from baseline will be summarized descriptively by treatment approach and by visit up to the end

TABLE 2 Analysis sets

Randomized analysis set (RAN) and full analysis set (FAS)	RAN comprises all randomized patients and FAS comprises all randomized patients who received at least one dose of randomized study medication (vildagliptin or placebo) and have at least one post-randomization assessment of any efficacy parameter.
Intent-to-treat analysis (ITT)	Following the ITT principle, data for participants will be analysed according to the treatment approach to which they were assigned at randomization. Patients will not be excluded from ITT analysis based on protocol deviations, including violations of study entry criteria concerning prior use of anti-diabetic agents and, eg, BMI range.
Safety set (SAF)	For assessment of safety, the SAF comprises all patients who received at least one dose of randomized study medication (vildagliptin or placebo). Patients will be analysed according to the treatment approach received. If a patient would have erroneously received, eg, both vildagliptin and placebo during Period 1, the patient will be included in the vildagliptin group. NOTE: the SAF allows inclusion of non-randomized patients who received the study drug in error, or those who did not tolerate up-titration of metformin and were thus excluded from the study prior to randomization.
Per protocol set (PPS)	PPS is a subset of FAS and comprises all randomized patients who received at least one dose of randomized study medication (vildagliptin or placebo), who have undergone at least one post-randomization assessment of any efficacy parameter during Period 1, who did not discontinue the study prior to week 26 (earliest time-point for assessment of primary endpoint with consecutive, confirmed measurement of initial loss of glycaemic control) and for whom there were no major protocol deviations as assessed prior to database lock during Period 1.
Screened-only set (SCR)	SCR comprises all patients who were screen-failed after the first visit or who entered the run-in phase but were not randomized. No other analysis will be performed on this analysis set.

The number and percentage of patients in each analysis set will be summarized by treatment approach. The number and percentage of patients included in the subgroup analyses, which required a separate consent at selected sites, concerning an opportunity to join the meal test, retinal micro-aneurysm count and biomarker subsets (European patients only), will also be summarized by treatment approach. Meal test and retinal micro-aneurysm count data will be presented as part of the efficacy evaluations. No summaries or analyses will be created for biomarker data as part of this analysis plan, as the samples, drawn after receipt of a separate consent, are part of a pivotal consortium initiative for assessment of drug-induced liver toxicity in a European population.

of Periods 1 and 2 and the end of study. Additionally, descriptive sub-group analyses will be performed for primary and key secondary analysis based on pre-defined patient characteristics and other baseline co-variates (Table 3). Other subgroups may be considered as needed.

3.6 | Exploratory analyses

Exploratory endpoints include time to insulin initiation and change in body weight over time. HbA1c thresholds for failure and success will be examined. Homeostatic model assessment will be used to assess beta cell function (HOMA-B) and insulin resistance (HOMA-IR) will be calculated as a function of glucose and insulin, using both the original formula[20] and the interactive iHOMA2[21] model. Implementation of this model in the VERIFY study will be used to evaluate the impact of study treatment strategies and to predict their effects on fasting glucose, insulin, β-cell function and insulin sensitivity.

Changes in health status, determined using the EQ-5D questionnaire,[22] will be assessed in conjunction with the selected adverse events (AEs) and will be included in various health economic modelling plans, especially for local assessment of the cost effectiveness of combination therapy as compared to a stepwise approach.

TABLE 3 Pre-planned sub-group analyses

Sub-group co-variates	Definitions for analysis cut-off values
Baseline HbA1c category	<7%, ≥7%
Baseline BMI	<30 kg/m^2, ≥30 kg/m^2
Age	per baseline tertiles of <51 years (n = 626), ≥59 years (n = 743)
Gender	Male / female
Baseline smoking status	Yes / no
Race	White Caucasians, Asians, Hispanic Americans, etc.
Geographical regions	Europe, Asia (with and without India), Latin America, Australia, South Africa
Associations between beta cell function and insulin resistance	Based on assessment of HOMA-IR, HOMA-β / iHOMA-%β, iHOMA-% sensitivity
Other baseline demographics	As appropriate

Abbreviations: BMI, body mass index; HbA1c, glycated haemoglobin; HOMA-IR, homeostasis model assessments for insulin resistance; HOMA-β, homeostatic model assessment-%β; HOMA-β, homeostasis model assessments for β-cell function; HOMA% sensitivity homeostatic model assessment % sensitivity.

In a subgroup of patients, change in retinal micro-aneurysm count from baseline to Years 4 and 5 will be assessed by independent, central reading, based on a pre-defined protocol. Microvascular and macrovascular complications will be descriptively reported, including summaries of microalbuminuria, progression to renal insufficiency and all-cause mortality. For secondary and exploratory endpoints, the time-to-event variables will be analysed using a Cox proportional hazard model similar to that used for the primary variable; the rate of loss in respective variables will be analysed using a linear mixed effect model; the variables for change from baseline will be evaluated by an analysis of covariance (ANCOVA), with treatment approach and geographic region as classification factors and baseline as a covariate.

3.7 | Avoiding bias

Pre-planned protection against risk of bias or confounding has been essential in the design of the VERIFY study. Mitigation of and attempts to minimize bias is achieved by optimization of the study design through randomization and blinding, through inclusion/exclusion criteria to avoid selection bias, through a robust retention plan to detect bias for true effect or outcomes, and through intention-to-treat analysis. Public submission of a detailed SAP is a key component to avoid occurrence of *post-hoc* data selection or selective reporting and, ultimately, publication bias.

The VERIFY study included patients who reflective the current, rather variable T2DM management standards. In order to avoid measurement or analytical bias, all samples were tranferred for analysis to an accredited central laboratory, to confirm eligibility, and also to determine all key variables. The eligibility criteria ensured inclusion of a wide range of newly diagnosed patients, ensuring that their glycaemic parameters were within the anticipated range and that changes in glycaemia, the main endpoint, was not influenced by concomitant disease(s) or management thereof. Initially, based on previous studies in newly diagnosed patients,[19] albeit in a less diverse population, the rate of premature discontinuation was anticipated to be significantly higher. Considerable efforts have been made to maximize the completeness of study participation via an effective retention programme, to avoid under-estimation of the true effects of treatment strategies on most outcomes, while the likely impact can be quantified and the direction of the effect is known, for the most part. However, sensitivity analyses will be performed as well. For example, the first occurrence of the initial loss of glycaemic control will be used to eliminate the bias induced by knowledge of the progressive glycaemic change or by the impact of clinical reality. It must be acknowledged that longitudinal changes, in body weight for example, might introduce an unintentional confounding effect. Therefore, a robust analysis of the role of the independent determinants that drive the primary response will be undertaken as part of a post hoc data interrogation, in addition to this pre-planned SAP. Results will be reported in accordance with CONSORT guidelines for cohort studies[23] and will be submitted for publication in a peer-reviewed journal.

4 | *FOCUS* ON SAFETY AND EXPLORATION OF ADJUDICATING CARDIOVASCULAR ENDPOINTS

An integrated analysis of safety and tolerability will compare the two treatment strategies during the five-year treatment periods in the safety data set (SAF). Key safety variables (overall AEs, serious AEs, AEs leading to study drug discontinuation or interruption, incidence of hypoglycaemia, predefined AE risks of interest) will also be summarized by treatment approach and expectations (Table 4). SAEs will be grouped into major categories, defined by the latest Medical Dictionary for Regulatory Activities version 21.1 (MedDRA).[24] The incidence of treatment-emergent AEs will be summarized by primary system organ class, preferred term, severity and relationship to study drug. Evaluation of clinical laboratory test results and vital signs will be undertaken only at scheduled study visits, while any laboratory abnormalities necessitating treatment or meeting a clinically significant change will be captured on the database as an AE and included in AE summary tables/listings.

TABLE 4 Safety assessments

Category
1. Physical examination
2. Vital signs
3. Body weight
4. Haematology: RBC (total), WBC (total), platelet count (direct), haemoglobin, haematocrit, basophils (absolute, %), eosinophils (absolute, %), lymphocytes (absolute, %), monocytes (absolute, %), neutrophils (absolute, %)
5. Blood chemistry: ALT, albumin, alkaline phosphatase, AST, bilirubin (direct), bilirubin (total), blood urea nitrogen, calcium (total), chloride, high-sensitivity c-reactive protein, phosphocreatine kinase (CPK), creatinine kinase muscle-brain-type isoenzyme if CPK elevated, creatinine, γ-glutamyl transferase, GFR via MDRD formula, lactate dehydrogenase, potassium, protein (total), sodium, uric acid
6. Urine tests: blood, glucose, ketones, leukocytes, pH, protein, pregnancy (β-hCG), urine albumin/creatinine ratio
7. Liver function test: ALT, AST, bilirubin (direct), bilirubin (total), alkaline phosphatase
8. Electrocardiogram
9. Pregnancy
10. Incidence of hypoglycaemia events
11. All treatment-emergent AEs/SAEs
12. New or progression of existing microvascular and macrovascular complications (reported as AEs), new onset microalbuminuria, progression to renal insufficiency (eGFR <60 mL min^{-1} 1.73 m^{-2}) or doubling of serum creatinine to at least 200 μM (2.26 mg/dL) and all-cause mortality, to assess the overall complications event rate during the study

Abbreviations: AE, adverse events; ALT, Alanine aminotransferase; AST, aspartate aminotransferase; eGFR, estimated glomerular filtration rate; hCG, human chorionic gonadotropin; MDRD, modification of diet in renal disease; RBC, red blood cells; SAE, serious AE; WBC, white blood cells.

Hypoglycaemic events (HEs) (<3.0 mmol/L [<54 mg/dL])[25] will be characterized by event profile, such as ability to self-treat and self-monitor plasma glucose level, precipitating event, time from last meal, time from last dose and time of day. The incidence of HEs and asymptomatic low blood glucose that was entered in the glycaemia study diary and met the criteria for those reporting at least one HE and for those who discontinued following HEs or who reported confirmed or suspected grade 2 events will be summarized by numbers and percentages for each treatment approach. The summary will be repeated according to pre-defined age groups and any other cut-offs as deemed appropriate after unblinding. Kaplan–Meier analyses will be used if the numbers are sufficient to warrant the approach.

Overall microvascular and macrovascular complications will be assessed separately by treatment approach. Overall complications are defined as new or progression of existing microvascular and macrovascular complications, confirmed by adjudication, new-onset microalbuminuria, progression to renal insufficiency (eGFR <60 mL min^{-1} 1.73 m^{-2}) or doubling of serum creatinine to at least 200 μM (2.26 mg/dL) and all-cause mortality. The frequency and percentage of patients experiencing each complication will be summarized.

5 | CHANGES TO THE ORIGINAL SAP

For long-term interventional studies, the SAP must be critically reviewed and updated as necessary prior to unblinding and data analysis. The reason for updates, such as baseline characteristics that drove the pre-planned sub-analyses in a previously unpredictable direction, will be clearly documented in the SAP to ensure transparency and external validity of the study.[26]

The original SAP for the VERIFY study was created based on the initial study protocol and related planned analyses, and it was updated/refined prior to completion of data collection and database lock. Key changes to the study protocol and/or SAP since the first version included additional clarification concerning a contraindication to treatment with metformin or vildagliptin. At the same time, a new inclusion criterion was added, based on health authority recommendations to ensure that only patients who received appropriate advice concerning lifestyle modification prior to enrolment, including diet counselling and exercise training, may be included. However, the main amendment to the protocol involved inclusion of a CV adjudication committee and an independent data monitoring committee. In the context of that amendment, additional dose-adjustment schedules and the addition of serum C-peptide measurement as a standard chemistry assessment were introduced for all patients.

Since the study commenced, it was decided to explore the blinded baseline characteristics as a *post-hoc* analysis as this was not prespecified in the study protocol. Additionally, the SAP was refined to reflect an updated understanding of the main analysis approach by introducing only one primary efficacy endpoint. As the probability and risk of time to initial loss of glycaemic control as the primary efficacy variable will be presented as a hazard ratio and failure rate over time, the statistical approach and power calculations, including the impact of the known retention, were updated accordingly, while assessment of the rate of loss of glycaemic control by an annualized slope of mean HbA1c over time from Week 26 to the end of Period 1 was designated, rather, as a key secondary analysis. Based on the now known baseline characteristics, the originally suggested sub-analyses have been modified to reflect the actual population with respect to distribution of age, weight and BMI categories.

6 | CONCLUSIONS

According to optimum trial practice, details of the statistical analysis and data handling plan prior to locking the VERIFY study database are reported here. This plan is coherent with the high quality standards of internal validity to minimize analysis bias, and will enhance the clinical utility of reported results, aiming to improve early treatment and leading to more optimized outcomes in the management of patients with newly diagnosed T2DM.

ACKNOWLEDGEMENTS

The authors thank Ranjan Tiwari and Qing Shao for current and initial statistical support, respectively; Plamen Kozlovski, Wolfgang Kothny and James Foley for clinical and scientific expertise and input at the study design phase; and Nihal Maremanda for providing editorial assistance.

CONFLICT OF INTEREST

D. R. M. has served on advisory boards or as a consultant for Novo Nordisk, GlaxoSmithKline, Novartis, Eli Lilly, Sanofi-Aventis, Janssen and Servier; receives current committee-related support from Janssen; has given lectures for Novo Nordisk, Servier, Sanofi-Aventis, Eli Lilly, Novartis, Janssen and Aché Laboratories; currently serves as President of the EASD. M. S. has received speaker's honoraria and consulting fees from Novartis, Novo Nordisk, AstraZeneca, Aegerion, Eli Lilly and Company and Boehringer Ingelheim. J. H., G. B., Y. C., P. M. P. and P. P. are employed by and own stocks in Novartis. S. D. P. serves or has served on advisory boards for AstraZeneca, Boehringer Ingelheim, Eli Lilly and Company, GlaxoSmithKline, Hanmi Pharmaceuticals, Intarcia, Janssen Pharmaceutics, Merck Sharp & Dohme Ltd, Novartis, Novo Nordisk, Sanofi, Servier and Takeda; serves or has served on tspeakers' bureaus for AstraZeneca, Boehringer Ingelheim, Eli Lilly and Company, Janssen Pharmaceutics, Merck Sharp & Dohme Ltd, Novartis, Novo Nordisk, Sanofi and Takeda; and has received research support from Boehringer Ingelheim, Merck Sharp & Dohme Ltd and Novartis.

AUTHOR CONTRIBUTIONS

The manuscript was drafted by P. M. P. The original protocol-defined assumptions and input for statistical considerations were provided by D. R. M., M. S. and S. D. P. and, based on this input, P. M. P. and P. P.

prepared the original SAP. All authors critically revised and contributed to the content of the updated SAP and this manuscript and approved the final submitted version.

ORCID

David R. Matthews https://orcid.org/0000-0001-6504-0036
Stefano Del Prato https://orcid.org/0000-0002-5388-0270

REFERENCES

1. King P, Peacock I, Donnelly R. The UKprospective diabetes study (UKPDS): clinical and therapeutic implications for type 2 diabetes. *Br J Clin Pharmacol*. 1999;48:643-648.
2. Holman RR, Paul SK, Bethel MA, Matthews DR, Neil HA. 10-year follow-up of intensive glucose control in type 2 diabetes. *N Engl J Med*. 2008;359:1577-1589.
3. Brath H, Paldánius PM, Bader G, Kolaczynski WM, Nilsson PM. Differences in glycemic control across world regions: a post-hoc analysis in patients with type 2 diabetes mellitus on dual antidiabetes drug therapy. *Nutr Diabetes*. 2016;6:e217.
4. Singh AK, Singh R. SAVOR-TIMI to SUSTAIN-6: a critical comparison of cardiovascular outcome trials of antidiabetic drugs. *Expert Rev Clin Pharmacol*. 2017;10:429-442.
5. Al-Arouj M, Hassoun AA, Medlej R, et al. The effect of vildagliptin relative to sulphonylureas in Muslim patients with type 2 diabetes fasting during Ramadan: the VIRTUE study. *Int J Clin Pract*. 2013;67:957-963.
6. Kozlovski P, Fonseca M, Mohan V, et al. Effect of race and ethnicity on vildagliptin efficacy: A pooled analysis of phase II and III studies. *Diabetes Obes Metab*. 2017;19:429-435.
7. Strain WD, Lukashevich V, Kothny W, Hoellinger MJ, Paldánius PM. Individualised treatment targets for elderly patients with type 2 diabetes using vildagliptin add-on or lone therapy (INTERVAL): a 24 week, randomised, double-blind, placebo-controlled study. *Lancet*. 2013;382:409-416.
8. Mathieu C, Barnett AH, Brath H, et al. Effectiveness and tolerability of second-line therapy with vildagliptin vs. other oral agents in type 2 diabetes: a real-life worldwide observational study (EDGE). *Int J Clin Pract*. 2013;67:947-956.
9. Davies MJ, D'Alessio DA, Fradkin J, et al. Management of hyperglycaemia in type 2 diabetes, 2018. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetologia*. 2018;61:2461-2498.
10. Del Prato S, Felton AM, Munro N, et al; Global Partnership for Effective Diabetes Management.Improving glucose management: ten steps to get more patients with type 2 diabetes to glycaemic goal. *Int J Clin Pract*. 2005;59:1345-1355.
11. Del Prato S, Foley JE, Kothny W, et al. Study to determine the durability of glycaemic control with early treatment with a vildagliptin-metformin combination regimen vs. standard-of-care metformin monotherapy-the VERIFY trial: a randomised double-blind trial. *Diabet Med*. 2014;31:1178-1184.
12. Scherbaum WA, Schweizer A, Mari A, et al. Evidence that vildagliptin attenuates deterioration of glycaemic control during 2-year treatment of patients with type 2 diabetes and mild hyperglycaemia. *Diabetes Obes Metab*. 2008;10:1114-1124.
13. Cefalu WT, Kaul S, Gerstein HC, et al. Cardiovascular outcomes trials in type 2 diabetes: where do we go from here? Reflections from a diabetes care editors' expert forum. *Diabetes Care*. 2018;41:14-31.
14. Scirica BM, Bhatt DL, Braunwald E, et al; SAVOR-TIMI 53 Steering Committee and Investigators.Saxagliptin and cardiovascular outcomes in patients with type 2 diabetes mellitus. *N Engl J Med*. 2013;369:1317-1326.
15. Mathieu C, Kozlovski P, Paldánius PM, et al. Clinical safety and tolerability of vildagliptin - insights from randomised trials, observational studies and post-marketing surveillance. *Eur Endocrinol*. 2017;13:68-72.
16. Matthews DR, Paldánius PM, Proot P, Foley JE, Stumvoll M, del Prato S. Baseline characteristics in the VERIFY study: a randomised trial assessing the durability of glycaemic control with early vildagliptin-metformin combination in newly diagnosed type 2 diabetes. *Diabet Med*. 2019;36:505-513.
17. Raman M, Middleton RJ, Kalra PA, Green D. Estimating renal function in old people: an in-depth review. *Int Urol Nephrol*. 2017;49:1979-1988.
18. Nathan DM, Buse JB, Davidson MB, et al. Medical management of hyperglycemia in type 2 diabetes: a consensus algorithm for the initiation and adjustment of therapy: a consensus statement of the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care*. 2009;32:193-203.
19. Kahn SE, Haffner SM, Heise MA, et al; ADOPT Study Group.Glycemic durability of rosiglitazone, metformin, or glyburide monotherapy. *N Engl J Med*. 2006;355:2427-2443.
20. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*. 1985;28:412-419.
21. Hill NR, Levy JC, Matthews DR. Expansion of the homeostasis model assessment of β-cell function and insulin resistance to enable clinical trial outcome modeling through the interactive adjustment of physiology and treatment effects: iHOMA2. *Diabetes Care*. 2013;36:2324-2330.
22. Group TE. EuroQol-a new facility for the measurement of health-related quality of life. *Health policy*. 1990;16:199-208.
23. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c869.
24. Brown EG, Wood L, Wood S. The medical dictionary for regulatory activities (MedDRA). *Drug Saf*. 1999;20:109-117.
25. International Hypoglycaemia Study Group. Glucose concentrations of less than 3.0 mmol/L (54 mg/dL) should be reported in clinical trials: a joint position statement of the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care*. 2017;40:155-157.
26. Gamble C, Krishan A, Stocken D, et al. Guidelines for the content of statistical analysis plans in clinical trials. *JAMA*. 2017;318:2337-2343.

How to cite this article: Matthews DR, Paldánius PM, Stumvoll M, et al. A pre-specified statistical analysis plan for the VERIFY study: Vildagliptin efficacy in combination with metformin for early treatment of T2DM. *Diabetes Obes Metab*. 2019;21:2240–2247. https://doi.org/10.1111/dom.13800