D2. Network Meta-Analysis Supplemental Information

NMA Methods

We evaluated the feasibility of conducting a quantitative synthesis by exploring the differences in study populations, study design, analytic methods, and outcome assessment for each outcome of interest. Trials deemed sufficiently similar in terms of population, intervention type, duration, and outcome definitions were included in the NMA. Based on data availability, we developed quantitative, indirect comparisons of tirzepatide, injectable semaglutide, empagliflozin, and background therapy using a Bayesian NMA for outcomes of change in HbA1c, weight, LDL, and SBP at 40 weeks in adult patients with T2DM. The primary endpoints of the tirzepatide trial, SURPASS-2 was measured at 40 weeks, PIONEER-2 and 3 were measured at 52 weeks, SUSTAIN-2 at 56 weeks, and HARMONY-3 at 104 weeks. We received academic-in-confidence outcomes data at week 40 from manufacturers of four of the five trials in the NMA. We were unable to access week 40 data from the HARMONY-3 trial. Thus, we used digitized estimates from published figures to calculate change in HbA1c and weight at week 40; we were limited to using change from baseline to week 104 for the LDL and SBP outcomes. For the outcomes of HbA1c and weight, results were reported using the treatment-regimen or equivalent estimand. For the outcome of LDL, results were reported using the efficacy estimand for SURPASS-2 and treatment-regimen estimand equivalent for SUSTAIN-2, PIONEER-2 and 3.

All four outcomes were analyzed as continuous outcomes using a generalized linear model with identity link (Tables D2.1-2.3). Vague priors were used for study-specific baselines and basic parameters in the Bayesian NMA models. However, between-study heterogeneity could not be precisely estimated in the random effect models with noninformative priors due to the small number of studies available. This was reflected by the wide credible intervals of the summary effects (data available upon request). We considered applying fixed-effect models; however, the fixed-effect model does not account for variation in intervention effects across studies. Therefore, we used random-effects models for all four outcomes and made assumptions about the extent of heterogeneity. We used the information provided by Rhodes et al. to construct informative priors for the between-study variance (τ^2).[1] (see Table D2.1) All NMAs were conducted using the IndiRect NMA platform (CRG-EVERSANA, 2020™). We initially discarded the first 10,000 iterations as "burn-in" and based inferences on an additional 50,000 iterations using three chains. We evaluated the convergence of chains through visual examination of the Brook–Gelman–Rubin diagnostic and historical plots.

Table D2.1. NMAs Conducted & Presented

Outcome	Model	Number of trials
Change from Baseline in HbA1c (%)	Generalized linear model with identity link; Random effect model with informative prior (prior for between study variance: log t[−3.68,2.78^2,5])[1]	5
Change from Baseline in Body Weight (kg)	Generalized linear model with identity link; Random effect with informative prior (prior for between study variance: log t[−3.44,2.44^2,5])[1]	
Change from Baseline in LDL (mg/dL)	Generalized linear model with identity link; Random effect model with informative prior (prior for between study variance: log t[−3.68,2.78^2,5])[1]	
Change from Baseline in SBP (mmHg)	Generalized linear model with identity link; Random effect with informative prior (prior for between study variance: log t[−3.44,2.44^2,5])[1]	

%: percentage point, HbA1c: hemoglobin A1c, kg: kilogram, LDL: low-density lipoprotein, mg/dL: milligram/deciliter, mmHg: millimeters of mercury, SBP: systolic blood pressure

Table D2.2. Data Inputs for NMA of HbA1c and Body Weight Loss

		HbA1C (%)			Body Weight (kg)		
Trial	**Name**	**N**	**Mean**	**Standard Error**	**N**	**Mean**	**Standard Error**
SURPASS-2	TZP		-2.3			-11.2	
	SEM		-1.86			-5.7	
SUSTAIN 2	SEM						
	SITA						
PIONEER 2	OSEM						
	EMPA						
PIONEER 3	OSEM						
	SITA						
HARMONY 3*	PBO	101	0	0.13	101	2.2	1
	SITA	302	-0.5	0.07	302	0.4	0.95

%: percentage point, EMPA: empagliflozin, HBA1c: hemoglobin A1c, kg: kilogram, N: number, OSEM: oral semaglutide, PBO: placebo, SEM: semaglutide, SITA: sitagliptin, TZP: tirzepatide

*Values from HARMONY 3 were digitized by ICER staff

Table D2.3. Data Inputs for NMA of LDL and SBP

		LDL Cholesterol (mg/dL)			Systolic Blood Pressure (mmHg)		
Trial	**Name**	**N**	**Mean**	**Standard Error**	**N**	**Mean**	**Standard Error**
SURPASS-2	TZP	[redacted]	-4.5	[redacted]	[redacted]	-6.5	[redacted]
	SEM		-5.6			-3.6	
SUSTAIN 2	SEM						
	SITA						
PIONEER 2	OSEM						
	EMPA						
PIONEER 3	OSEM						
	SITA						
HARMONY 3*	PBO	101	-1.2	2.7	101	2.2	1.39
	SITA	302	-1.9	1.34	302	0.2	0.85

EMPA: empagliflozin, LDL: low-density lipoprotein, mg/dL: milligrams per deciliter, N: number, OSEM: oral semaglutide, PBO: placebo, SBP: systolic blood pressure, SEM: semaglutide, SITA: sitagliptin, TZP: tirzepatide
*Note: For LDL and SBP, mean values of change were calculated from baseline to week 104

Table D2.4. NMA Results for Change in HbA1c, Body Weight, LDL and SBP at Week 40 Between Tirzepatide versus Background Therapy/ Empagliflozin

	TZP vs BT	**TZP vs EMPA**
	Mean Difference (95% Credible Interval)	
Change in HbA1c, %,	-1.72 (-1.93 to -1.51)	-1.12 (-1.37 to -0.88)
Change in Body Weight, kg	-11.51 (-12.13 to -10.87)	-7.24 (-7.96 to -6.53)
Change in LDL, mg/dL	-4.34 (-5.48 to -3.23)	-7.53 (-8.72 to -6.33)
Change in SBP, mmHg	-7.46 (-8.16 to -6.76)	-2.59 (-3.37 to -1.82)

%: percentage point, BT: background therapy, EMPA: empagliflozin, HbA1c: hemoglobin A1c, kg: kilogram, LDL: low-density lipoprotein, mg/dL: milligram per deciliter, mmHg: millimeters of mercury, SBP: systolic blood pressure, TZP: tirzepatide

References

1. Rhodes KM, Turner RM, Higgins JP. Predictive distributions were developed for the extent of heterogeneity in meta-analyses of continuous outcome data. *Journal of clinical epidemiology.* 2015;68(1):52-60.