Hierarchical network meta-analysis models for synthesis of evidence from randomised and non-randomised studies

Phimphone Visalath

M.Sc Medical epidemiology

Hussein H, Abrams KR, Gray LJ, Anwer S, Dias S, Bujkiewicz S. Hierarchical network meta-analysis models for synthesis of evidence from randomised and non-randomised studies. BMC Med Res Methodol. 2023 Apr 22;23(1):97

Previous studies

- Schmitz et al. (2012)
- Method: naïve pooling, use of informative prior distributions and hierarchical models
- Rheumatoid arthritis.
- Result: inclusion of observational evidence in NMA increased uncertainty of the pooled effectiveness estimates.

Previous studies

- Jenkins et al. (2021)
- Method: naïve pooling, hierarchical model and power prior analysis
- Relapsing remitting multiple sclerosis
- Result: increased uncertainty compared to the analysis of RCT data alone, due to the increased between-study heterogeneity when incorporating data from non-randomised studies.

Introduction

- Hussein et al. (2019)
- Patient with type 2 diabetes
- Two classes of glucose-lowering medications:

1. sodium-glucose co-transporter 2 inhibitors (SGLT-2is)
2. glucagon-like peptide-1 receptor agonists (GLP-1RAs)

- Outcome: Glycated haemoglobin
- Study design: Network meta-analysis of

1. Randomised controlled trials (RCTs)
2. Non-RCTs

Method

- Naïve Pooling
- Hierarchical Models
- Bias Adjustment Models

To assessing their impact on the effect estimates and uncertainty.

Method

- ***Network meta-analysis models for inclusion of non-randomised data***
- ***Model A – naïve pooling***
- The basic NMA model was applied to both RCT and non-randomised data combined, with no adjustments made for different sources of data or classes of treatments within the network.

Method

- ***Model B1 – two-level hierarchical model (treatment vs class)***
- A two-level hierarchical model with treatments nested within treatment classes; i.e. treatments were nested within either SGLT-2i, GLP-1RA or placebo classes.
- The model allows for borrowing of information across treatments within each class when estimating pooled treatment effects for individual treatments, which are of primary interest.
- It also allows for estimating an average effect within each treatment class, which may also be of interest.

Method

- ***Model B2 – two-level hierarchical model (treatment vs design)***
- Modelling the between-study heterogeneity of treatment effects within and across each study design (i.e. RCT and non-randomised studies).
- The model allows for differentiating between treatment effects from studies of different designs when estimating pooled treatment effects for individual treatments, and it allows for estimation of these average effects for each type of study design individually, and also overall across all studies (whilst taking into account of the across-design heterogeneity).

Method

- ***Model B3 – three-level hierarchical model***
- This model was developed to extend the above two level models (B1 and B2), by allowing for an additional level in the random-effects hierarchical NMA model to estimate the heterogeneity within study designs as well as estimating the heterogeneity within treatment classes in the network.

Method

- ***Model C1 – bias adjustment assuming same bias by class***
- Observational studies are assumed to have additional risk of bias due to the absence of randomisation and unmeasured confounding.
- The bias adjustment model allows for this by including an additional bias parameter.
- different non-randomised studies vary in terms of the level of bias assumed.

Method

- ***Model C2 – bias adjustment assuming varying bias by class***
- The above bias adjustment model assumes exchangeable biases across all observational studies regardless of treatment class being compared.
- The level of bias may differ across classes.

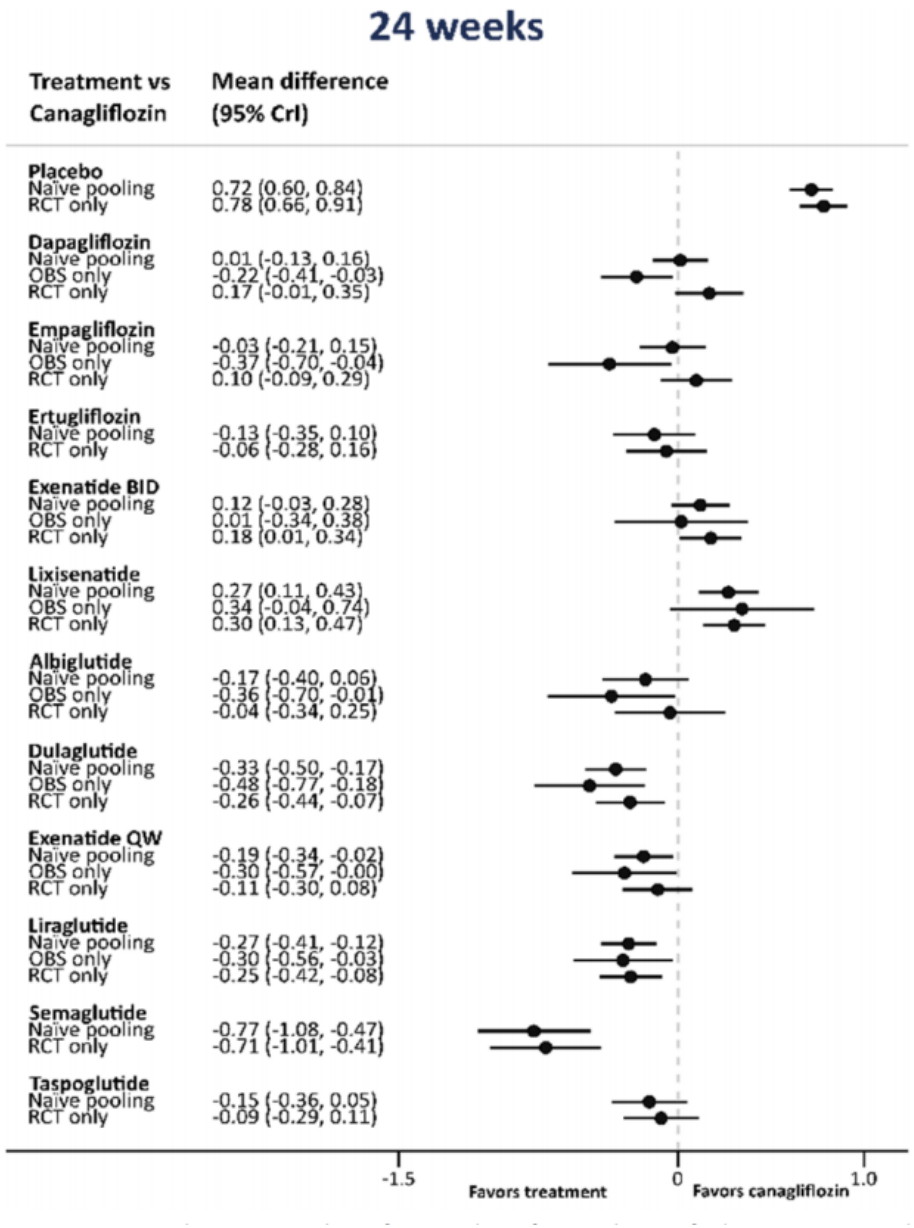
Result

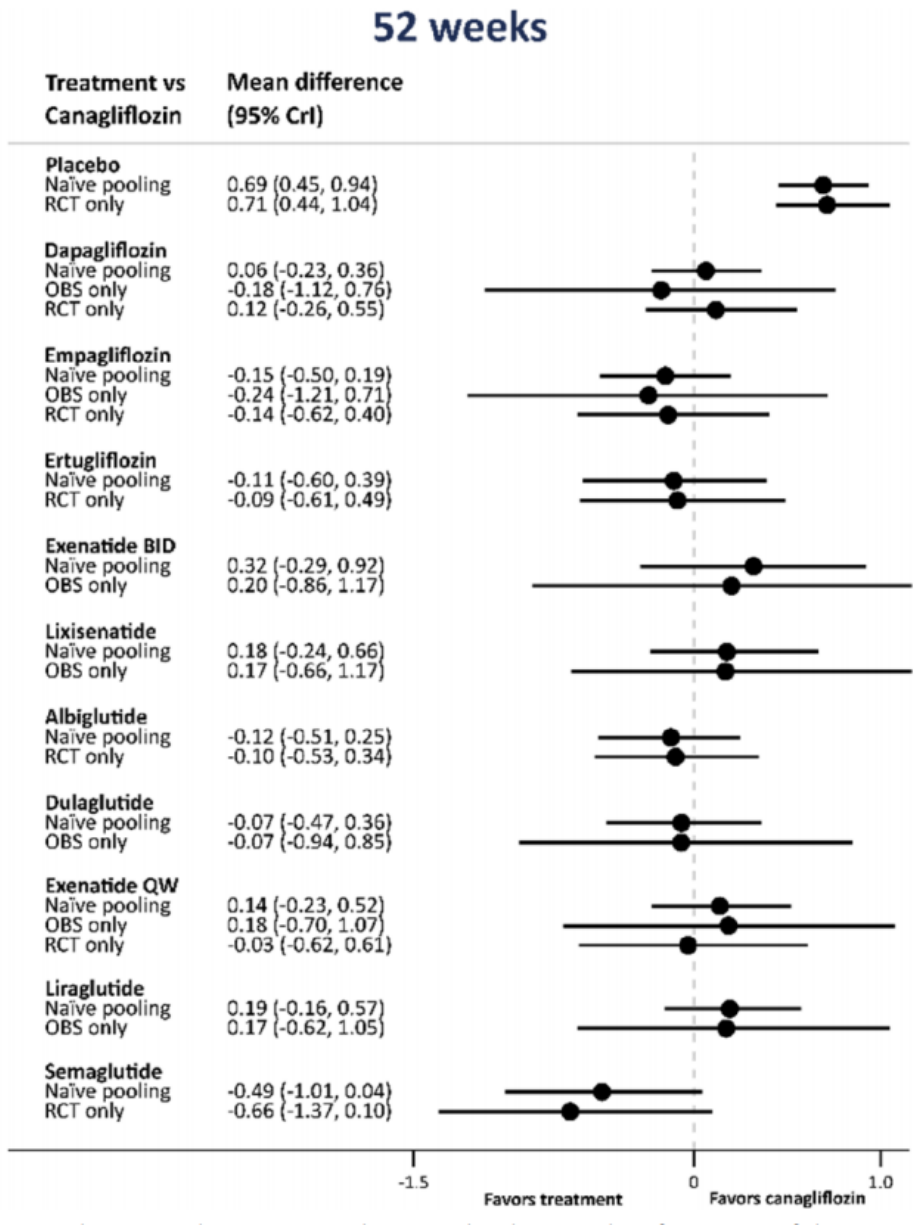
- 74 studies were included in this NMA.
- 64 papers were RCTs and 10 studies were non-randomised.

- The number of individuals with type 2 diabetes recruited to
  - RCTs on average 490 individuals (range: 50–2072 individuals)
  - Observational studies on average studying larger populations (mean: 1863, range: 212–5141 individuals).

Result

- Naïve pooling





Result

- Mean difference from baseline for HbA1c (%) at 24 weeks for all models fitted vs reference treatment (canaglifozin)

Treatment vs Canagliflozin	Model							
	RCT only	OBS only	Model A	Model B1	Model B2	Model B3	Model C1	Model C2
Effectiveness								
Dapagliflozin[a]	0.17 (-0.01, 0.35)	-0.22 (-0.41, -0.03)	0.01 (-0.13, 0.16)	-0.00 (-0.14, 0.14)	0.00 (-0.24, 0.23)	-0.02 (-0.22, 0.19)	0.04 (-0.12, 0.22)	0.10 (-0.05, 0.25)
Empagliflozin[a]	0.10 (-0.09, 0.30)	-0.37 (-0.70, -0.04)	-0.03 (-0.21, 0.15)	-0.03 (-0.19, 0.13)	-0.05 (-0.34, 0.20)	-0.05 (-0.28, 0.16)	0.00 (-0.19, 0.19)	0.03 (-0.13, 0.19)
Exenatide BID	0.18 (0.01, 0.35)	0.01 (-0.34, 0.38)	0.12 (-0.04, 0.28)	0.11 (-0.04, 0.27)	0.13 (-0.13, 0.39)	0.09 (-0.16, 0.32)	0.13 (-0.02, 0.29)	0.12 (-0.02, 0.26)
Lixisenatide	0.30 (0.13, 0.47)	0.34 (-0.04, 0.74)	0.27 (0.11, 0.43)	0.26 (0.1, 0.42)	0.31 (0.06, 0.59)	0.24 (-0.06, 0.48)	0.27 (0.11, 0.43)	0.25 (0.11, 0.39)
Albiglutide	-0.04 (-0.34, 0.25)	-0.36 (-0.70, -0.01)	-0.17 (-0.40, 0.06)	-0.17 (-0.39, 0.05)	-0.16 (-0.45, 0.13)	-0.15 (-0.41, 0.10)	-0.14 (-0.38, 0.10)	-0.13 (-0.34, 0.07)
Dulaglutide	-0.26 (-0.44, -0.07)	-0.48 (-0.77, -0.18)	-0.33 (-0.50, -0.17)	-0.33 (-0.49, -0.17)	-0.32 (-0.58, -0.07)	-0.3 (-0.52, -0.06)	-0.31 (-0.48, -0.15)	-0.33 (-0.47, -0.18)
Exenatide QW	-0.11 (-0.30, 0.08)	-0.30 (-0.57, 0.00)	-0.19 (-0.34, -0.02)	-0.19 (-0.34, -0.03)	-0.16 (-0.41, 0.09)	-0.16 (-0.37, 0.06)	-0.17 (-0.33, 0.00)	-0.18 (-0.32, -0.03)
Liraglutide	-0.25 (-0.42, -0.08)	-0.30 (-0.56, -0.03)	-0.27 (-0.41, -0.12)	-0.26 (-0.41, -0.12)	-0.24 (-0.48, 0.00)	-0.23 (-0.44, 0.01)	-0.28 (-0.43, -0.12)	-0.31 (-0.45, -0.18)
Placebo	0.78 (0.66, 0.91)	-	0.72 (0.60, 0.84)	0.72 (0.61, 0.84)	0.74 (0.44, 1.06)	0.74 (0.47, 1.08)	0.73 (0.61, 0.85)	0.73 (0.63, 0.83)
Ertugliflozin[a]	-0.06 (-0.28, 0.16)	-	-0.13 (-0.35, 0.09)	-0.08 (-0.3, 0.1)	-0.10 (-0.44, 0.25)	-0.07 (-0.34, 0.18)	-0.11 (-0.33, 0.10)	-0.11 (-0.30, 0.08)
Semaglutide	-0.71 (-1.01, -0.41)	-	-0.77 (-1.08, -0.47)	-0.67 (-0.97, -0.37)	-0.75 (-1.14, -0.35)	-0.58 (-0.95, -0.14)	-0.76 (-1.06, -0.46)	-0.76 (-1.04, -0.48)
Taspoglutide	-0.09 (-0.29, 0.11)	-	-0.15 (-0.36, 0.05)	-0.15 (-0.35, 0.04)	-0.13 (-0.46, 0.22)	-0.13 (-0.4, 0.14)	-0.14 (-0.34, 0.06)	-0.14 (-0.32, 0.04)
Class-level effects and bias								
D. SGLT-2i				-0.04 (-0.55, 0.46)		-0.05 (-0.66, 0.55)		
D. GLP-1RA				-0.17 (-0.48, 0.12)		-0.15 (-0.45, 0.14)		
Bias							0.06 (-0.09, 0.2)	
Bias. SGLT-2i								-0.24 (-0.43, -0.03)
Bias. GLP-1RA								0.13 (0.00, 0.25)
Heterogeneity								
SD	0.1 (0.04, 0.16)	0.04 (0.00, 0.17)	0.11 (0.07, 0.16)	0.11 (0.07, 0.16)	0.09 (0.03, 0.14)	0.09 (0.02, 0.14)	0.1 (0.05, 0.16)	0.08 (0.01, 0.13)
SD. SGLT-2i				0.1 (0.0, 1.78)		0.12 (0.0, 2.03)		
SD. GLP-1RA				0.33 (0.18, 0.7)		0.31 (0.12, 0.69)		
SD.design					0.11 (0.02, 0.28)	0.1 (0.02, 0.26)		
SD.bias							0.09 (0.0, 0.26)	0.04 (0.0, 0.15)

Result

- Mean difference from baseline for HbA1c (%) at 52 weeks for all models fitted vs reference treatment (canaglifozin)

Treatment vs Canagliflozin	Model							
	RCT only	OBS only	Model A	Model B1	Model B2	Model B3	Model C1	Model C2
Effectiveness								
Dapagliflozin[a]	0.12 (-0.26, 0.55)	-0.18 (-1.12, 0.76)	0.06 (-0.23, 0.37)	0.03 (-0.27, 0.33)	0.01 (-0.72, 0.68)	-0.02 (-0.39, 0.34)	0.10 (-0.18, 0.40)	0.06 (-0.22, 0.34)
Empagliflozin[a]	-0.14 (-0.62, 0.40)	-0.24 (-1.21, 0.50)	-0.15 (-0.51, 0.20)	-0.12 (-0.47, 0.21)	-0.17 (-0.90, 0.53)	-0.14 (-0.53, 0.24)	-0.15 (-0.48, 0.20)	-0.15 (-0.45, 0.17)
Exenatide BID	-	0.20 (-0.86, 0.53)	0.32 (-0.28, 0.90)	0.14 (-0.32, 0.62)	0.27 (-0.84, 1.30)	0.08 (-0.4, 0.6)	0.49 (-0.46, 1.30)	0.14 (-0.80, 1.07)
Lixisenatide	-	0.17 (-0.66, 0.48)	0.18 (-0.24, 0.67)	0.1 (-0.28, 0.62)	0.14 (-0.89, 1.14)	0.04 (-0.4, 0.5)	0.24 (-0.40, 0.88)	-0.11 (-0.82, 0.66)
Dulaglutide	-	-0.07 (-0.94, 0.46)	-0.07 (-0.47, 0.36)	-0.05 (-0.42, 0.34)	-0.07 (-1.07, 0.93)	-0.06 (-0.49, 0.37)	0.05 (-0.66, 0.66)	-0.20 (-0.91, 0.48)
Exenatide QW	-0.03 (-0.61, 0.60)	0.18 (-0.70, 0.46)	0.14 (-0.23, 0.52)	0.14 (-0.42, 0.34)	0.10 (-0.64, 0.81)	0.08 (-0.32, 0.47)	0.10 (-0.30, 0.51)	-0.09 (-0.54, 0.37)
Liraglutide	-	0.17 (-0.62, 0.43)	0.19 (-0.16, 0.57)	0.14 (-0.21, 0.49)	0.18 (-0.83, 1.15)	0.07 (-0.36, 0.5)	0.08 (-0.49, 0.72)	-0.25 (-0.89, 0.47)
Placebo	0.71 (0.44, 1.04)	-	0.69 (0.45, 0.94)	0.71 (0.47, 0.96)	0.70 (-0.23, 1.66)	0.72 (0.28, 1.17)	0.70 (0.49, 0.94)	0.69 (0.49, 0.90)
Ertugliflozin[a]	-0.09 (-0.61, 0.48)	-	-0.11 (-0.60, 0.40)	-0.07 (-0.5, 0.35)	-0.10 (-1.10, 0.92)	-0.08 (-0.55, 0.39)	-0.10 (-0.53, 0.36)	-0.11 (-0.52, 0.31)
Albiglutide	-0.10 (-0.53, 0.34)	-	-0.12 (-0.52, 0.25)	-0.07 (-0.4, 0.31)	-0.11 (-1.07, 0.87)	-0.06 (-0.44, 0.35)	-0.11 (-0.45, 0.23)	-0.12 (-0.44, 0.19)
Semaglutide	-0.66 (-1.36, 0.09)	-	-0.49 (-1.01, 0.05)	-0.29 (-0.76, 0.26)	-0.59 (-1.69, 0.42)	-0.25 (-0.8, 0.28)	-0.54 (-1.03, 0.00)	-0.72 (-1.26, 0.18)
Class-level effects and bias								
d. SGLT-2i				-0.05 (-0.99, 0.85)		-0.08 (-1.05, 0.87)		
d. GLP-1RA				0.02 (-0.36, 0.41)		-0.01 (-0.4, 0.39)		
Bias							0.21 (-0.5, 0.8)	
Bias. SGLT-2i								-0.21 (-1.06, 0.61)
Bias. GLP-1RA								0.26 (-0.34, 0.78)
Heterogeneity								
SD	0.10 (0.01, 0.5)	0.27 (0.02, 1.04)	0.13 (0.01, 0.33)	0.14 (0.01, 0.33)	0.11 (0.01, 0.32)	0.11 (0.01, 0.3)	0.09 (0.00, 0.3)	0.08 (0.00, 0.28)
SD. SGLT-2i				0.23 (0.01, 2.7)		0.23 (0.01, 2.81)		
SD. GLP-1RA				0.23 (0.03, 0.64)		0.2 (0.01, 0.64)		
SD.design					0.21 (0.01, 1.38)	0.14 (0.01, 0.39)a		
SD.bias							0.19 (0.01, 0.62)	0.14 (0.01, 0.57)

- Naïve-pooling averaged the effect estimate between what was observed in RCTs and non-randomised studies, with most effect estimates having similar or smaller credible intervals in comparison to the results of NMA of RCT data alone.
- Hierarchical models fitted accounted for the design of the study, which was further extended to consider the classification of treatments within the SGLT-2i and GLP-1RA class, effect estimates were similar to those from the naïve pooling method but credible intervals were often wider.

Conclusion

- The inclusion of observational data in NMAs of RCTs is gaining considerable traction in HTA due to the many benefits:

1. Increasing evidence base
2. Potentially connecting disconnected networks
3. Allowing for more generalizable inferences.

- Both, hierarchical and bias adjustment models can provide a better fit to the data in comparison to naïve pooling and should be explored when conducting evidence synthesis.

Thank you