Supplementary material

Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in a meta-analysis of randomized controlled trials

Version 1.1

Consider the following important instructions informed by common misapplications of ICEMAN in studies using the instrument

- Complete a separate credibility assessment per each effect modifier (e.g., age, comorbidity, drug dose, etc.), outcome (e.g., mortality, stroke, duration of hospital stay), time-point (e.g., 3 months, 6 months), and effect measure (e.g. relative risk, risk difference).
- Do not apply ICEMAN if the interaction p-value is 0.1 or larger, i.e., provides very little statistical support for the existence of an effect modification (ICEMAN is designed to address the possible claim of an effect modification rather than the claim of no effect modification).
- Response options on the left indicate definitely or probably reduced credibility, response options on the right probably or definitely increased credibility
- Completely unclear should be interpreted as probably reduced credibility.
- To ensure transparency, provide a supporting comment under each question that provides a rationale for the rating.
- To ensure transparency, provide a copy of the completed ICEMAN instrument in the supplement of your article.

CREDIBILITY ASSESSMENT

Essential preliminary considerations to define the possible effect modification of interest

State a single candidate effect modifier (e.g., age or comorbidity): Co-treatment of sodium-glucose co-transporter-2 inhibitors (SGLT-2I)/glucagon-like peptide-1 receptor agonists (GLP-1RA)

Was the effect modifier measured before or at randomization? [ ] yes, continue [**X**] no, stop here and refer to manual for further instructions

State a single outcome and time-point (e.g., mortality at 1 year follow-up): Major adverse cardiovascular events (MACE) within the follow-up period of major cardiovascular outcomes trials (CVOT)

State a single effect measure (e.g., relative risk or risk difference): Hazard ratio

1: Is the analysis of effect modification based on comparison within rather than between trials?

[**X**] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Subgroup analysis or meta-regression comparing overall effects of each individual trial. This is typical for aggregate data meta-analysis.	*Subgroup analysis or meta-regression with most information coming from overall effects, but some trials providing within-trial subgroup information*	*Most trials providing within-trial subgroup information; or individual participant data analysis that combines within and between trial information*	*All trials providing within-trial subgroup information or individual participant data; and the analysis separates within from between trial information, e.g., meta-analysis of interactions*

Comment: This network meta-regression incorporated the major CVOT to investigate potential cardiovascular benefit of combination treatment of GLP-1RA and SGLT-2I in patients with type 2 diabetes (T2D).

2: For within-trial comparisons, is the effect modification similar from trial to trial? [ ] Not applicable: no or one within-RCT comparison

[ ] Definitely not similar	[ ] Probably not similar or unclear	[ ] Mostly similar	[**X**] Definitely similar
Effect modification reported for two or more trials and clearly different directions	*Effect modification not reported for individual trials or too imprecise to tell*	*Effect modification reported for two or more trials, mostly similar in direction, but considerable differences in magnitude*	*Effect modification reported for two or more trials, similar in direction, only some differences in magnitude*

Comment: Several post hoc analyses of GLP-1RA/SGLT-2I have demonstrated that the combination treatment of GLP-1RA and SGLT-2I can reduce MACE in patients with T2D.

3: For between-trial comparisons, is the number of trials large? [ ] Not applicable: no between RCT comparison

[ ] Very small	[**X**] Rather small or unclear	[ ] Rather large	[ ] Large
1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 14 CVOT of GLP-1RA and SGLT-2I (including 1 post hoc analysis) were included in the network meta-regression. However, the data points are still considered scarce, and the regression model might be over specified.

4: Was the direction of effect modification correctly hypothesized a priori?

[ ] Definitely no	[ ] Probably no or unclear	[**X**] Probably yes	[ ] Definitely yes

Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of effect modification*	*Prior protocol available and includes correct specification of direction of effect modification, e.g., based on a biologic rationale*

Comment: We correctly anticipated the cardiovascular benefit of the combination treatment of GLP-1RA and SGLT-2I. However, this hypothesis was not included in the protocol published on PROSPERO.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification? (consider irrespective of number of effect modifiers)

[ ] Chance a very likely explanation	[**X**] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The mean and 95% CI of the coefficient (β) [-0.07 (-0.22, 0.08)] suggest that the correlation – effect modification of the co-treatment on MACE may be negative.

6: Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[**X**] Definitely yes
Explicitly exploratory analysis or large number of effect modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 effect modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer effect modifiers tested*	*Protocol available and 3 or fewer effect modifiers tested or number considered in analysis*

Comment: Only effect modifications of the co-treatment of GLP-RA/SGLT-2I were assessed on MACE in patients with T2D. The protocol was available in PROSPERO.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[**X**] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: A random effect Bayesian model was applied in the network meta-regression.

8: If the effect modifier is a continuous variable, were arbitrary cut points avoided? [**X**] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior RCT*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment:

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [ ] not applicable

[ ] Yes, probably decrease	[**X**] Yes, probably increase

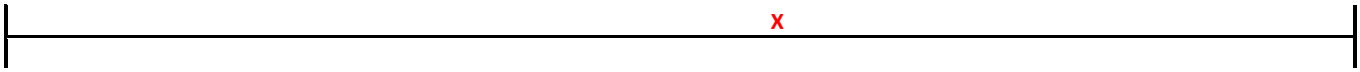
Comment: The cardiovascular benefit of the combination treatment of GLP-1RA and SGLT-2I are consistent across the different outcomes.

10: How would you rate the overall credibility of the proposed effect modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Minimal to no support for effect modification; Use overall effect for each subgroup	Some but insufficient support for effect modification; Use overall effect for each subgroup but note remaining uncertainty	Likely effect modification; Use separate effects for each subgroup but note remaining uncertainty	Very likely effect modification; Use separate effects for each subgroup

Comment: The regression model might become over specified due to scarcity of the data points. The β was generated with less statistical power (because of the over-specification of the model).

Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in a meta-analysis of randomized controlled trials

Version 1.1

Consider the following important instructions informed by common misapplications of ICEMAN in studies using the instrument

- Complete a separate credibility assessment per each effect modifier (e.g., age, comorbidity, drug dose, etc.), outcome (e.g., mortality, stroke, duration of hospital stay), time-point (e.g., 3 months, 6 months), and effect measure (e.g. relative risk, risk difference).
- Do not apply ICEMAN if the interaction p-value is 0.1 or larger, i.e., provides very little statistical support for the existence of an effect modification (ICEMAN is designed to address the possible claim of an effect modification rather than the claim of no effect modification).
- Response options on the left indicate definitely or probably reduced credibility, response options on the right probably or definitely increased credibility
- Completely unclear should be interpreted as probably reduced credibility.
- To ensure transparency, provide a supporting comment under each question that provides a rationale for the rating.
- To ensure transparency, provide a copy of the completed ICEMAN instrument in the supplement of your article.

CREDIBILITY ASSESSMENT

Essential preliminary considerations to define the possible effect modification of interest

State a single candidate effect modifier (e.g., age or comorbidity): Co-treatment of sodium-glucose co-transporter-2 inhibitors (SGLT-2I)/glucagon-like peptide-1 receptor agonists (GLP-1RA)

Was the effect modifier measured before or at randomization? [ ] yes, continue [X] no, stop here and refer to manual for further instructions

State a single outcome and time-point (e.g., mortality at 1 year follow-up): Cardiovascular death within the follow-up period of major cardiovascular outcomes trials (CVOT)

State a single effect measure (e.g., relative risk or risk difference): Hazard ratio

1: Is the analysis of effect modification based on comparison within rather than between trials?

[X] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Subgroup analysis or meta-regression comparing overall effects of each individual trial. This is typical for aggregate data meta-analysis.	*Subgroup analysis or meta-regression with most information coming from overall effects, but some trials providing within-trial subgroup information*	*Most trials providing within-trial subgroup information; or individual participant data analysis that combines within and between trial information*	*All trials providing within-trial subgroup information or individual participant data; and the analysis separates within from between trial information, e.g., meta-analysis of interactions*

Comment: This network meta-regression incorporated major CVOT to investigate potential cardiovascular benefit of combination treatment of GLP-1RA and SGLT-2I in patients with type 2 diabetes (T2D).

2: For within-trial comparisons, is the effect modification similar from trial to trial? [ ] Not applicable: no or one within-RCT comparison

[ ] Definitely not similar	[ ] Probably not similar or unclear	[ ] Mostly similar	[X] Definitely similar
Effect modification reported for two or more trials and clearly different directions	*Effect modification not reported for individual trials or too imprecise to tell*	*Effect modification reported for two or more trials, mostly similar in direction, but considerable differences in magnitude*	*Effect modification reported for two or more trials, similar in direction, only some differences in magnitude*

Comment: Several post hoc analyses of GLP-1RA/SGLT-2I have demonstrated that the combination treatment of GLP-1RA and SGLT-2I can reduce cardiovascular death in patients with T2D.

3: For between-trial comparisons, is the number of trials large? [ ] Not applicable: no between RCT comparison

[ ] Very small	[X] Rather small or unclear	[ ] Rather large	[ ] Large
1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 14 CVOT of GLP-1RA and SGLT-2I (including 1 post hoc analysis) were included in the network meta-regression. However, the data points are still considered scarce, and the regression model might be over specified.

4: Was the direction of effect modification correctly hypothesized a priori?

[ ] Definitely no	[ ] Probably no or unclear	[X] Probably yes	[ ] Definitely yes

Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of effect modification*	*Prior protocol available and includes correct specification of direction of effect modification, e.g., based on a biologic rationale*

Comment: We correctly anticipated the cardiovascular benefit of the combination treatment of GLP-1RA and SGLT-2I. However, this hypothesis was not included in the protocol published on PROSPERO.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification? (consider irrespective of number of effect modifiers)

[ ] Chance a very likely explanation	[**X**] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The mean and 95% CI of the coefficient (β) [-0.06 (-0.31, 0.17)] suggest that the correlation – effect modification of the co-treatment on cardiovascular death may be negative.

6: Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[**X**] Definitely yes
Explicitly exploratory analysis or large number of effect modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 effect modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer effect modifiers tested*	*Protocol available and 3 or fewer effect modifiers tested or number considered in analysis*

Comment: Only effect modifications of the co-treatment of GLP-RA/SGLT-2I and prior cardiovascular diseases were assessed on cardiovascular death in patients with T2D. The protocol was available in PROSPERO.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[**X**] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: A random effect Bayesian model was applied in the network meta-regression.

8: If the effect modifier is a continuous variable, were arbitrary cut points avoided? [**X**] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior RCT*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment:

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [ ] not applicable

[ ] Yes, probably decrease	[**X**] Yes, probably increase

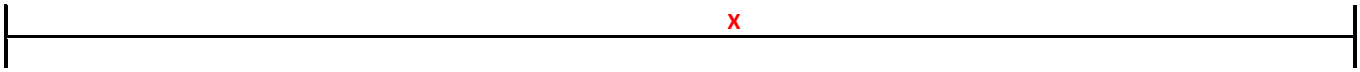
Comment: The cardiovascular benefit of the combination treatment of GLP-1RA and SGLT-2I are consistent across the different outcomes.

10: How would you rate the overall credibility of the proposed effect modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Minimal to no support for effect modification; Use overall effect for each subgroup	Some but insufficient support for effect modification; Use overall effect for each subgroup but note remaining uncertainty	Likely effect modification; Use separate effects for each subgroup but note remaining uncertainty	Very likely effect modification; Use separate effects for each subgroup

Comment: The regression model might become over specified due to scarcity of the data points. The β was generated with less statistical power (because of the over-specification of the model).

Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in a meta-analysis of randomized controlled trials

Version 1.1

Consider the following important instructions informed by common misapplications of ICEMAN in studies using the instrument

- Complete a separate credibility assessment per each effect modifier (e.g., age, comorbidity, drug dose, etc.), outcome (e.g., mortality, stroke, duration of hospital stay), time-point (e.g., 3 months, 6 months), and effect measure (e.g. relative risk, risk difference).
- Do not apply ICEMAN if the interaction p-value is 0.1 or larger, i.e., provides very little statistical support for the existence of an effect modification (ICEMAN is designed to address the possible claim of an effect modification rather than the claim of no effect modification).
- Response options on the left indicate definitely or probably reduced credibility, response options on the right probably or definitely increased credibility
- Completely unclear should be interpreted as probably reduced credibility.
- To ensure transparency, provide a supporting comment under each question that provides a rationale for the rating.
- To ensure transparency, provide a copy of the completed ICEMAN instrument in the supplement of your article.

CREDIBILITY ASSESSMENT

Essential preliminary considerations to define the possible effect modification of interest

State a single candidate effect modifier (e.g., age or comorbidity): Co-treatment of sodium-glucose co-transporter-2 inhibitors (SGLT-2I)/glucagon-like peptide-1 receptor agonists (GLP-1RA)

Was the effect modifier measured before or at randomization? [ ] yes, continue [X] no, stop here and refer to manual for further instructions

State a single outcome and time-point (e.g., mortality at 1 year follow-up): Fatal and non-fatal myocardial infraction (MI) within the follow-up period of major cardiovascular outcomes trials (CVOT)

State a single effect measure (e.g., relative risk or risk difference): Hazard ratio

1: Is the analysis of effect modification based on comparison within rather than between trials?

[X] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Subgroup analysis or meta-regression comparing overall effects of each individual trial. This is typical for aggregate data meta-analysis.	*Subgroup analysis or meta-regression with most information coming from overall effects, but some trials providing within-trial subgroup information*	*Most trials providing within-trial subgroup information; or individual participant data analysis that combines within and between trial information*	*All trials providing within-trial subgroup information or individual participant data; and the analysis separates within from between trial information, e.g., meta-analysis of interactions*

Comment: This network meta-regression incorporated major CVOT to investigate potential cardiovascular benefit of combination treatment of GLP-1RA and SGLT-2I in patients with type 2 diabetes (T2D).

2: For within-trial comparisons, is the effect modification similar from trial to trial? [ ] Not applicable: no or one within-RCT comparison

[ ] Definitely not similar	[X] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Effect modification reported for two or more trials and clearly different directions	*Effect modification not reported for individual trials or too imprecise to tell*	*Effect modification reported for two or more trials, mostly similar in direction, but considerable differences in magnitude*	*Effect modification reported for two or more trials, similar in direction, only some differences in magnitude*

Comment: Only a meta-analysis of HARMONY OUTCOMES and AMPLITUDE-O has demonstrated that the combination treatment of GLP-1RA and SGLT-2I can reduce the composite outcome of cardiovascular death, MI, unstable angina or all-cause mortality, MI, stroke.

3: For between-trial comparisons, is the number of trials large? [ ] Not applicable: no between RCT comparison

[ ] Very small	[X] Rather small or unclear	[ ] Rather large	[ ] Large
1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 14 CVOT of GLP-1RA and SGLT-2I (including 1 post hoc analysis) were included in the network meta-regression. However, the data points are still considered scarce, and the regression model might be over specified.

4: Was the direction of effect modification correctly hypothesized a priori?

[ ] Definitely no	[ ] Probably no or unclear	[X] Probably yes	[ ] Definitely yes

Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of effect modification*	*Prior protocol available and includes correct specification of direction of effect modification, e.g., based on a biologic rationale*

Comment: We correctly anticipated the cardiovascular benefit of the combination treatment of GLP-1RA and SGLT-2I. However, this hypothesis was not included in the protocol published on PROSPERO.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification? (consider irrespective of number of effect modifiers)

[ ] Chance a very likely explanation	[**X**] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The mean and 95% CI of the coefficient (β) [-0.09 (-0.43, 0.11)] suggest that the correlation – effect modification of the co-treatment on fatal and non-fatal MI may be negative.

6: Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[**X**] Definitely yes
Explicitly exploratory analysis or large number of effect modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 effect modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer effect modifiers tested*	*Protocol available and 3 or fewer effect modifiers tested or number considered in analysis*

Comment: Only effect modifications of the co-treatment of GLP-RA/SGLT-2I and prior cardiovascular diseases were assessed on fatal and non-fatal MI in patients with T2D. The protocol was available in PROSPERO.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[**X**] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: A random effect Bayesian model was applied in the network meta-regression.

8: If the effect modifier is a continuous variable, were arbitrary cut points avoided? [**X**] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior RCT*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment:

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [ ] not applicable

[ ] Yes, probably decrease	[**X**] Yes, probably increase

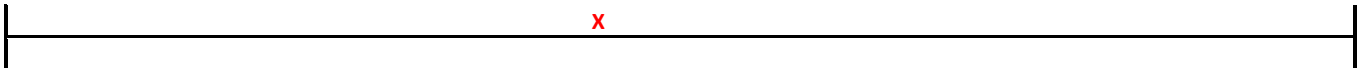
Comment: The cardiovascular benefit of the combination treatment of GLP-1RA and SGLT-2I are consistent across the different outcomes.

10: How would you rate the overall credibility of the proposed effect modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Minimal to no support for effect modification; Use overall effect for each subgroup	Some but insufficient support for effect modification; Use overall effect for each subgroup but note remaining uncertainty	Likely effect modification; Use separate effects for each subgroup but note remaining uncertainty	Very likely effect modification; Use separate effects for each subgroup

Comment: Consistency across studies was unclear. The regression model might become over specified due to scarcity of the data points. The β was generated with less statistical power (because of the over-specification of the model).

Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in a meta-analysis of randomized controlled trials

Version 1.1

Consider the following important instructions informed by common misapplications of ICEMAN in studies using the instrument

- Complete a separate credibility assessment per each effect modifier (e.g., age, comorbidity, drug dose, etc.), outcome (e.g., mortality, stroke, duration of hospital stay), time-point (e.g., 3 months, 6 months), and effect measure (e.g. relative risk, risk difference).
- Do not apply ICEMAN if the interaction p-value is 0.1 or larger, i.e., provides very little statistical support for the existence of an effect modification (ICEMAN is designed to address the possible claim of an effect modification rather than the claim of no effect modification).
- Response options on the left indicate definitely or probably reduced credibility, response options on the right probably or definitely increased credibility
- Completely unclear should be interpreted as probably reduced credibility.
- To ensure transparency, provide a supporting comment under each question that provides a rationale for the rating.
- To ensure transparency, provide a copy of the completed ICEMAN instrument in the supplement of your article.

CREDIBILITY ASSESSMENT

Essential preliminary considerations to define the possible effect modification of interest

State a single candidate effect modifier (e.g., age or comorbidity): Co-treatment of sodium-glucose co-transporter-2 inhibitors (SGLT-2I)/glucagon-like peptide-1 receptor agonists (GLP-1RA)

Was the effect modifier measured before or at randomization? [ ] yes, continue [X] no, stop here and refer to manual for further instructions

State a single outcome and time-point (e.g., mortality at 1 year follow-up): Fatal and non-fatal stroke within the follow-up period of major cardiovascular outcomes trials (CVOT)

State a single effect measure (e.g., relative risk or risk difference): Hazard ratio

1: Is the analysis of effect modification based on comparison within rather than between trials?

[X] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Subgroup analysis or meta-regression comparing overall effects of each individual trial. This is typical for aggregate data meta-analysis.	*Subgroup analysis or meta-regression with most information coming from overall effects, but some trials providing within-trial subgroup information*	*Most trials providing within-trial subgroup information; or individual participant data analysis that combines within and between trial information*	*All trials providing within-trial subgroup information or individual participant data; and the analysis separates within from between trial information, e.g., meta-analysis of interactions*

Comment: This network meta-regression incorporated major CVOT to investigate potential cardiovascular benefit of combination treatment of GLP-1RA and SGLT-2I in patients with type 2 diabetes (T2D).

2: For within-trial comparisons, is the effect modification similar from trial to trial? [ ] Not applicable: no or one within-RCT comparison

[ ] Definitely not similar	[X] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Effect modification reported for two or more trials and clearly different directions	*Effect modification not reported for individual trials or too imprecise to tell*	*Effect modification reported for two or more trials, mostly similar in direction, but considerable differences in magnitude*	*Effect modification reported for two or more trials, similar in direction, only some differences in magnitude*

Comment: Only a meta-analysis of HARMONY OUTCOMES and AMPLITUDE-O has demonstrated that the combination treatment of GLP-1RA and SGLT-2I can reduce the composite outcome of cardiovascular death, MI, unstable angina or all-cause mortality, MI, stroke.

3: For between-trial comparisons, is the number of trials large? [ ] Not applicable: no between RCT comparison

[ ] Very small	[X] Rather small or unclear	[ ] Rather large	[ ] Large
1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 14 CVOT of GLP-1RA and SGLT-2I (including 1 post hoc analysis) were included in the network meta-regression. However, the data points are still considered scarce, and the regression model might be over specified.

4: Was the direction of effect modification correctly hypothesized a priori?

[ ] Definitely no	[ ] Probably no or unclear	[X] Probably yes	[ ] Definitely yes

Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of effect modification*	*Prior protocol available and includes correct specification of direction of effect modification, e.g., based on a biologic rationale*

Comment: We correctly anticipated the cardiovascular benefit of the combination treatment of GLP-1RA and SGLT-2I. However, this hypothesis was not included in the protocol published on PROSPERO.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification? (consider irrespective of number of effect modifiers)

[**X**] Chance a very likely explanation	[ ] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The mean and 95% CI of the coefficient (β) [-0.01 (-0.26, 0.24)] suggest that the correlation – effect modification of the co-treatment on fatal and non-fatal stroke may be neutral.

6: Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[**X**] Definitely yes
Explicitly exploratory analysis or large number of effect modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 effect modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer effect modifiers tested*	*Protocol available and 3 or fewer effect modifiers tested or number considered in analysis*

Comment: Only effect modifications of the co-treatment of GLP-RA/SGLT-2I were assessed on fatal and non-fatal stroke in patients with T2D. The protocol was available in PROSPERO.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[**X**] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: A random effect Bayesian model was applied in the network meta-regression.

8: If the effect modifier is a continuous variable, were arbitrary cut points avoided? [**X**] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior RCT*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment:

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [ ] not applicable

[ ] Yes, probably decrease	[**X**] Yes, probably increase

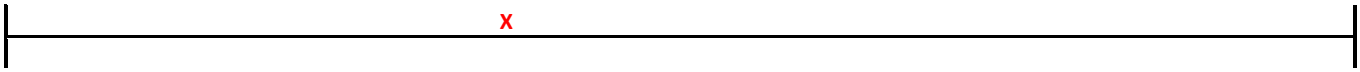
Comment: The cardiovascular benefit of the combination treatment of GLP-1RA and SGLT-2I are consistent across the different outcomes.

10: How would you rate the overall credibility of the proposed effect modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Minimal to no support for effect modification; Use overall effect for each subgroup	Some but insufficient support for effect modification; Use overall effect for each subgroup but note remaining uncertainty	Likely effect modification; Use separate effects for each subgroup but note remaining uncertainty	Very likely effect modification; Use separate effects for each subgroup

Comment: Consistency across studies was unclear. The regression model might become over specified due to scarcity of the data points. The β was generated with less statistical power (because of the over-specification of the model).

Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in a meta-analysis of randomized controlled trials

Version 1.1

Consider the following important instructions informed by common misapplications of ICEMAN in studies using the instrument

- Complete a separate credibility assessment per each effect modifier (e.g., age, comorbidity, drug dose, etc.), outcome (e.g., mortality, stroke, duration of hospital stay), time-point (e.g., 3 months, 6 months), and effect measure (e.g. relative risk, risk difference).
- Do not apply ICEMAN if the interaction p-value is 0.1 or larger, i.e., provides very little statistical support for the existence of an effect modification (ICEMAN is designed to address the possible claim of an effect modification rather than the claim of no effect modification).
- Response options on the left indicate definitely or probably reduced credibility, response options on the right probably or definitely increased credibility
- Completely unclear should be interpreted as probably reduced credibility.
- To ensure transparency, provide a supporting comment under each question that provides a rationale for the rating.
- To ensure transparency, provide a copy of the completed ICEMAN instrument in the supplement of your article.

CREDIBILITY ASSESSMENT

Essential preliminary considerations to define the possible effect modification of interest

State a single candidate effect modifier (e.g., age or comorbidity): Co-treatment of sodium-glucose co-transporter-2 inhibitors (SGLT-2I)/glucagon-like peptide-1 receptor agonists (GLP-1RA)

Was the effect modifier measured before or at randomization? [ ] yes, continue [X] no, stop here and refer to manual for further instructions

State a single outcome and time-point (e.g., mortality at 1 year follow-up): Hospitalization for heart failure (HF) within the follow-up period of major cardiovascular outcomes trials (CVOT)

State a single effect measure (e.g., relative risk or risk difference): Hazard ratio

1: Is the analysis of effect modification based on comparison within rather than between trials?

[X] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Subgroup analysis or meta-regression comparing overall effects of each individual trial. This is typical for aggregate data meta-analysis.	*Subgroup analysis or meta-regression with most information coming from overall effects, but some trials providing within-trial subgroup information*	*Most trials providing within-trial subgroup information; or individual participant data analysis that combines within and between trial information*	*All trials providing within-trial subgroup information or individual participant data; and the analysis separates within from between trial information, e.g., meta-analysis of interactions*

Comment: This network meta-regression incorporated major CVOT to investigate potential cardiovascular benefit of combination treatment of GLP-1RA and SGLT-2I in patients with type 2 diabetes (T2D).

2: For within-trial comparisons, is the effect modification similar from trial to trial? [ ] Not applicable: no or one within-RCT comparison

[ ] Definitely not similar	[ ] Probably not similar or unclear	[ ] Mostly similar	[X] Definitely similar
Effect modification reported for two or more trials and clearly different directions	*Effect modification not reported for individual trials or too imprecise to tell*	*Effect modification reported for two or more trials, mostly similar in direction, but considerable differences in magnitude*	*Effect modification reported for two or more trials, similar in direction, only some differences in magnitude*

Comment: Several post hoc analyses of GLP-1RA/SGLT-2I have demonstrated that the combination treatment of GLP-1RA and SGLT-2I can reduce hospitalization for HF in patients with T2D.

3: For between-trial comparisons, is the number of trials large? [ ] Not applicable: no between RCT comparison

[ ] Very small	[X] Rather small or unclear	[ ] Rather large	[ ] Large
1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 14 CVOT of GLP-1RA and SGLT-2I (including 1 post hoc analysis) were included in the network meta-regression. However, the data points are still considered scarce, and the regression model might be over specified.

4: Was the direction of effect modification correctly hypothesized a priori?

[ ] Definitely no	[ ] Probably no or unclear	[X] Probably yes	[ ] Definitely yes

Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of effect modification*	*Prior protocol available and includes correct specification of direction of effect modification, e.g., based on a biologic rationale*

Comment: We correctly anticipated the cardiovascular benefit of the combination treatment of GLP-1RA and SGLT-2I. However, this hypothesis was not included in the protocol published on PROSPERO.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification? (consider irrespective of number of effect modifiers)

[ ] Chance a very likely explanation	[ ] Chance a likely explanation or unclear	[**X**] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The mean and 95% CI of the coefficient (β) [-0.13 (-0.42, 0.13)] suggest that the correlation – effect modification of the co-treatment on hospitalization for HF may be negative.

6: Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[**X**] Probably no or unclear	[ ] Probably yes	[**X**] Definitely yes
Explicitly exploratory analysis or large number of effect modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 effect modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer effect modifiers tested*	*Protocol available and 3 or fewer effect modifiers tested or number considered in analysis*

Comment: Only effect modifications of the co-treatment of GLP-RA/SGLT-2I and prior cardiovascular diseases were assessed on hospitalization for HF in patients with T2D. The protocol was available in PROSPERO.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[**X**] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: A random effect Bayesian model was applied in the network meta-regression.

8: If the effect modifier is a continuous variable, were arbitrary cut points avoided? [**X**] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior RCT*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment:

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [ ] not applicable

[ ] Yes, probably decrease	[**X**] Yes, probably increase

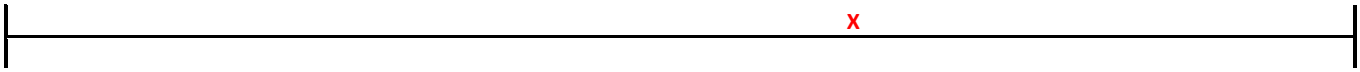
Comment: The cardiovascular benefit of the combination treatment of GLP-1RA and SGLT-2I are consistent across the different outcomes.

10: How would you rate the overall credibility of the proposed effect modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Minimal to no support for effect modification; Use overall effect for each subgroup	Some but insufficient support for effect modification; Use overall effect for each subgroup but note remaining uncertainty	Likely effect modification; Use separate effects for each subgroup but note remaining uncertainty	Very likely effect modification; Use separate effects for each subgroup

Comment: The regression model might become over specified due to scarcity of the data points. The β was generated with less statistical power (because of the over-specification of the model).

Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in a meta-analysis of randomized controlled trials

Version 1.1

Consider the following important instructions informed by common misapplications of ICEMAN in studies using the instrument

- Complete a separate credibility assessment per each effect modifier (e.g., age, comorbidity, drug dose, etc.), outcome (e.g., mortality, stroke, duration of hospital stay), time-point (e.g., 3 months, 6 months), and effect measure (e.g. relative risk, risk difference).
- Do not apply ICEMAN if the interaction p-value is 0.1 or larger, i.e., provides very little statistical support for the existence of an effect modification (ICEMAN is designed to address the possible claim of an effect modification rather than the claim of no effect modification).
- Response options on the left indicate definitely or probably reduced credibility, response options on the right probably or definitely increased credibility
- Completely unclear should be interpreted as probably reduced credibility.
- To ensure transparency, provide a supporting comment under each question that provides a rationale for the rating.
- To ensure transparency, provide a copy of the completed ICEMAN instrument in the supplement of your article.

CREDIBILITY ASSESSMENT

Essential preliminary considerations to define the possible effect modification of interest

State a single candidate effect modifier (e.g., age or comorbidity): Prior myocardial infarction (MI)

Was the effect modifier measured before or at randomization? [**X**] yes, continue [ ] no, stop here and refer to manual for further instructions

State a single outcome and time-point (e.g., mortality at 1 year follow-up): Fatal and non-fatal MI in patients with T2D receiving glucagon-like peptide-1 receptor agonists (GLP-1RA) within the follow-up period of major cardiovascular outcomes trials (CVOT)

State a single effect measure (e.g., relative risk or risk difference): Hazard ratio

1: Is the analysis of effect modification based on comparison within rather than between trials?

[**X**] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Subgroup analysis or meta-regression comparing overall effects of each individual trial. This is typical for aggregate data meta-analysis.	*Subgroup analysis or meta-regression with most information coming from overall effects, but some trials providing within-trial subgroup information*	*Most trials providing within-trial subgroup information; or individual participant data analysis that combines within and between trial information*	*All trials providing within-trial subgroup information or individual participant data; and the analysis separates within from between trial information, e.g., meta-analysis of interactions*

Comment: This network meta-regression incorporated major CVOT to investigate potential cardiovascular benefit of GLP-1RA in patients with type 2 diabetes (T2D).

2: For within-trial comparisons, is the effect modification similar from trial to trial? [ ] Not applicable: no or one within-RCT comparison

[ ] Definitely not similar	[**X**] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Effect modification reported for two or more trials and clearly different directions	*Effect modification not reported for individual trials or too imprecise to tell*	*Effect modification reported for two or more trials, mostly similar in direction, but considerable differences in magnitude*	*Effect modification reported for two or more trials, similar in direction, only some differences in magnitude*

Comment: There has not been any post hoc analysis investigating whether GLP-1RA can reduce fatal and non-fatal MI in patients with prior MI.

3: For between-trial comparisons, is the number of trials large? [ ] Not applicable: no between RCT comparison

[ ] Very small	[**X**] Rather small or unclear	[ ] Rather large	[ ] Large
1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 19 CVOT of GLP-1RA and SGLT-2I (including 6 post hoc analyses) were included in the network meta-regression. However, the data points are still considered scarce, and the regression model might be over specified.

4: Was the direction of effect modification correctly hypothesized a priori?

[ ] Definitely no	[**X**] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of effect modification*	*Prior protocol available and includes correct specification of direction of effect modification, e.g., based on a biologic rationale*

Comment: No information.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification? (consider irrespective of number of effect modifiers)

[ ] Chance a very likely explanation	[ ] Chance a likely explanation or unclear	[X] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The mean and 95% CI of the coefficient (β) [-0.17 (-0.40, 0.05)] suggest that the correlation – effect modification of prior MI on fatal and non-fatal MI in patients receiving GLP-1RA may be negative.

6: Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[X] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of effect modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 effect modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer effect modifiers tested*	*Protocol available and 3 or fewer effect modifiers tested or number considered in analysis*

Comment: Only effect modifications of the co-treatment of GLP-RA/SGLT-2I and prior cardiovascular diseases were assessed on fatal and non-fatal MI in patients with T2D, though the protocol is unavailable.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[X] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: A random effect Bayesian model was applied in the network meta-regression.

8: If the effect modifier is a continuous variable, were arbitrary cut points avoided? [X] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior RCT*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment:

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [X] not applicable

	[ ] Yes, probably decrease	[ ] Yes, probably increase	

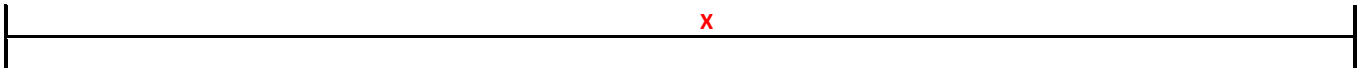
Comment:

10: How would you rate the overall credibility of the proposed effect modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Minimal to no support for effect modification; Use overall effect for each subgroup	Some but insufficient support for effect modification; Use overall effect for each subgroup but note remaining uncertainty	Likely effect modification; Use separate effects for each subgroup but note remaining uncertainty	Very likely effect modification; Use separate effects for each subgroup

Comment: Consistency across studies was unclear. Prior knowledge was unclear. The regression model might become over specified due to scarcity of the data points. The β was generated with less statistical power (because of the over-specification of the model).

Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in a meta-analysis of randomized controlled trials

Version 1.1

Consider the following important instructions informed by common misapplications of ICEMAN in studies using the instrument

- Complete a separate credibility assessment per each effect modifier (e.g., age, comorbidity, drug dose, etc.), outcome (e.g., mortality, stroke, duration of hospital stay), time-point (e.g., 3 months, 6 months), and effect measure (e.g. relative risk, risk difference).
- Do not apply ICEMAN if the interaction p-value is 0.1 or larger, i.e., provides very little statistical support for the existence of an effect modification (ICEMAN is designed to address the possible claim of an effect modification rather than the claim of no effect modification).
- Response options on the left indicate definitely or probably reduced credibility, response options on the right probably or definitely increased credibility
- Completely unclear should be interpreted as probably reduced credibility.
- To ensure transparency, provide a supporting comment under each question that provides a rationale for the rating.
- To ensure transparency, provide a copy of the completed ICEMAN instrument in the supplement of your article.

CREDIBILITY ASSESSMENT

Essential preliminary considerations to define the possible effect modification of interest

State a single candidate effect modifier (e.g., age or comorbidity): Prior myocardial infarction (MI)

Was the effect modifier measured before or at randomization? [**X**] yes, continue [ ] no, stop here and refer to manual for further instructions

State a single outcome and time-point (e.g., mortality at 1 year follow-up): Fatal and non-fatal MI in patients with T2D receiving SGLT-2I (sodium-glucose co-transporter-2 inhibitors) within the follow-up period of major cardiovascular outcomes trials (CVOT)

State a single effect measure (e.g., relative risk or risk difference): Hazard ratio

1: Is the analysis of effect modification based on comparison within rather than between trials?

[**X**] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Subgroup analysis or meta-regression comparing overall effects of each individual trial. This is typical for aggregate data meta-analysis.	*Subgroup analysis or meta-regression with most information coming from overall effects, but some trials providing within-trial subgroup information*	*Most trials providing within-trial subgroup information; or individual participant data analysis that combines within and between trial information*	*All trials providing within-trial subgroup information or individual participant data; and the analysis separates within from between trial information, e.g., meta-analysis of interactions*

Comment: This network meta-regression incorporated major CVOT to investigate potential cardiovascular benefit of SGLT-2I in patients with type 2 diabetes (T2D).

2: For within-trial comparisons, is the effect modification similar from trial to trial? [ ] Not applicable: no or one within-RCT comparison

[ ] Definitely not similar	[**X**] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Effect modification reported for two or more trials and clearly different directions	*Effect modification not reported for individual trials or too imprecise to tell*	*Effect modification reported for two or more trials, mostly similar in direction, but considerable differences in magnitude*	*Effect modification reported for two or more trials, similar in direction, only some differences in magnitude*

Comment: There has not been any post hoc analysis investigating whether SGLT-2I can reduce cardiovascular death in patients with prior MI

3: For between-trial comparisons, is the number of trials large? [ ] Not applicable: no between RCT comparison

[ ] Very small	[**X**] Rather small or unclear	[ ] Rather large	[ ] Large
1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 19 CVOT of GLP-1RA and SGLT-2I (including 6 post hoc analyses) were included in the network meta-regression. However, the data points are still considered scarce, and the regression model might be over specified.

4: Was the direction of effect modification correctly hypothesized a priori?

[ ] Definitely no	[**X**] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of effect modification*	*Prior protocol available and includes correct specification of direction of effect modification, e.g., based on a biologic rationale*

Comment: No information.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification? (consider irrespective of number of effect modifiers)

[ ] Chance a very likely explanation	[X] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The mean and 95% CI of the coefficient (β) [0.06 (-0.11, 0.24)] suggest that the correlation – effect modification of prior MI on fatal and non-fatal MI in patients receiving SGLT-2I may be positive.

6: Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[X] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of effect modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 effect modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer effect modifiers tested*	*Protocol available and 3 or fewer effect modifiers tested or number considered in analysis*

Comment: Only effect modifications of the co-treatment of GLP-RA/SGLT-2I and prior cardiovascular diseases were assessed on fatal and non-fatal MI in patients with T2D, though the protocol is unavailable.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[X] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: A random effect Bayesian model was applied in the network meta-regression.

8: If the effect modifier is a continuous variable, were arbitrary cut points avoided? [X] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior RCT*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment:

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [X] not applicable

	[ ] Yes, probably decrease	[ ] Yes, probably increase	

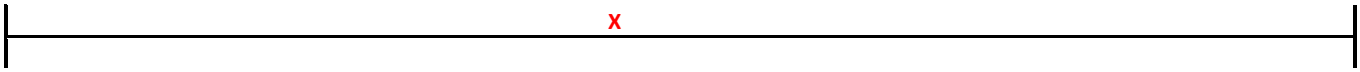
Comment: The cardiovascular benefit of the combination treatment of GLP-1RA and SGLT-2I are not consistent across the different outcomes.

10: How would you rate the overall credibility of the proposed effect modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Minimal to no support for effect modification; Use overall effect for each subgroup	Some but insufficient support for effect modification; Use overall effect for each subgroup but note remaining uncertainty	Likely effect modification; Use separate effects for each subgroup but note remaining uncertainty	Very likely effect modification; Use separate effects for each subgroup

Comment: Consistency across studies was unclear. Prior knowledge was unclear. The regression model might become over specified due to scarcity of the data points. The β was generated with less statistical power (because of the over-specification of the model).

Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in a meta-analysis of randomized controlled trials

Version 1.1

Consider the following important instructions informed by common misapplications of ICEMAN in studies using the instrument

- Complete a separate credibility assessment per each effect modifier (e.g., age, comorbidity, drug dose, etc.), outcome (e.g., mortality, stroke, duration of hospital stay), time-point (e.g., 3 months, 6 months), and effect measure (e.g. relative risk, risk difference).
- Do not apply ICEMAN if the interaction p-value is 0.1 or larger, i.e., provides very little statistical support for the existence of an effect modification (ICEMAN is designed to address the possible claim of an effect modification rather than the claim of no effect modification).
- Response options on the left indicate definitely or probably reduced credibility, response options on the right probably or definitely increased credibility
- Completely unclear should be interpreted as probably reduced credibility.
- To ensure transparency, provide a supporting comment under each question that provides a rationale for the rating.
- To ensure transparency, provide a copy of the completed ICEMAN instrument in the supplement of your article.

CREDIBILITY ASSESSMENT

Essential preliminary considerations to define the possible effect modification of interest

State a single candidate effect modifier (e.g., age or comorbidity): Prior myocardial infarction (MI)

Was the effect modifier measured before or at randomization? [**X**] yes, continue [ ] no, stop here and refer to manual for further instructions

State a single outcome and time-point (e.g., mortality at 1 year follow-up): Cardiovascular death in patients with T2D receiving glucagon-like peptide-1 receptor agonists (GLP-1RA) within the follow-up period of major cardiovascular outcomes trials (CVOT)

State a single effect measure (e.g., relative risk or risk difference): Hazard ratio

1: Is the analysis of effect modification based on comparison within rather than between trials?

[**X**] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Subgroup analysis or meta-regression comparing overall effects of each individual trial. This is typical for aggregate data meta-analysis.	*Subgroup analysis or meta-regression with most information coming from overall effects, but some trials providing within-trial subgroup information*	*Most trials providing within-trial subgroup information; or individual participant data analysis that combines within and between trial information*	*All trials providing within-trial subgroup information or individual participant data; and the analysis separates within from between trial information, e.g., meta-analysis of interactions*

Comment: This network meta-regression incorporated major CVOT to investigate potential cardiovascular benefit of GLP-1RA in patients with type 2 diabetes (T2D).

2: For within-trial comparisons, is the effect modification similar from trial to trial? [ ] Not applicable: no or one within-RCT comparison

[ ] Definitely not similar	[**X**] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Effect modification reported for two or more trials and clearly different directions	*Effect modification not reported for individual trials or too imprecise to tell*	*Effect modification reported for two or more trials, mostly similar in direction, but considerable differences in magnitude*	*Effect modification reported for two or more trials, similar in direction, only some differences in magnitude*

Comment: There has not been any post hoc analysis investigating whether GLP-1RA can reduce cardiovascular death in patients with prior MI

3: For between-trial comparisons, is the number of trials large? [ ] Not applicable: no between RCT comparison

[ ] Very small	[**X**] Rather small or unclear	[ ] Rather large	[ ] Large
1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 19 CVOT of GLP-1RA and SGLT-2I (including 6 post hoc analyses) were included in the network meta-regression. However, the data points are still considered scarce, and the regression model might be over specified.

4: Was the direction of effect modification correctly hypothesized a priori?

[ ] Definitely no	[**X**] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of effect modification*	*Prior protocol available and includes correct specification of direction of effect modification, e.g., based on a biologic rationale*

Comment: No information.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification? (consider irrespective of number of effect modifiers)

[ ] Chance a very likely explanation	[X] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The mean and 95% CI of the coefficient (β) [-0.06 (-0.43, 0.29)] suggest that the correlation – effect modification of prior MI on cardiovascular death in patients receiving GLP-1RA may be negative.

6: Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[X] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of effect modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 effect modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer effect modifiers tested*	*Protocol available and 3 or fewer effect modifiers tested or number considered in analysis*

Comment: Only effect modifications of the co-treatment of GLP-RA/SGLT-2I and prior cardiovascular diseases were assessed on cardiovascular death in patients with T2D, though the protocol is unavailable.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[X] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: A random effect Bayesian model was applied in the network meta-regression.

8: If the effect modifier is a continuous variable, were arbitrary cut points avoided? [X] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior RCT*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment:

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [X] not applicable

	[ ] Yes, probably decrease	[ ] Yes, probably increase	

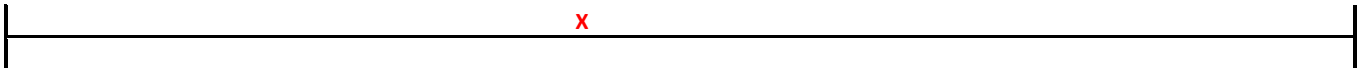
Comment:

10: How would you rate the overall credibility of the proposed effect modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Minimal to no support for effect modification; Use overall effect for each subgroup	Some but insufficient support for effect modification; Use overall effect for each subgroup but note remaining uncertainty	Likely effect modification; Use separate effects for each subgroup but note remaining uncertainty	Very likely effect modification; Use separate effects for each subgroup

Comment: Consistency across studies was unclear. Prior knowledge was unclear. The regression model might become over specified due to scarcity of the data points. The β was generated with less statistical power (because of the over-specification of the model).

Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in a meta-analysis of randomized controlled trials

Version 1.1

Consider the following important instructions informed by common misapplications of ICEMAN in studies using the instrument

- Complete a separate credibility assessment per each effect modifier (e.g., age, comorbidity, drug dose, etc.), outcome (e.g., mortality, stroke, duration of hospital stay), time-point (e.g., 3 months, 6 months), and effect measure (e.g. relative risk, risk difference).
- Do not apply ICEMAN if the interaction p-value is 0.1 or larger, i.e., provides very little statistical support for the existence of an effect modification (ICEMAN is designed to address the possible claim of an effect modification rather than the claim of no effect modification).
- Response options on the left indicate definitely or probably reduced credibility, response options on the right probably or definitely increased credibility
- Completely unclear should be interpreted as probably reduced credibility.
- To ensure transparency, provide a supporting comment under each question that provides a rationale for the rating.
- To ensure transparency, provide a copy of the completed ICEMAN instrument in the supplement of your article.

CREDIBILITY ASSESSMENT

Essential preliminary considerations to define the possible effect modification of interest

State a single candidate effect modifier (e.g., age or comorbidity): Prior myocardial infarction (MI)

Was the effect modifier measured before or at randomization? [**X**] yes, continue [ ] no, stop here and refer to manual for further instructions

State a single outcome and time-point (e.g., mortality at 1 year follow-up): Cardiovascular death in patients with T2D receiving SGLT-2I (sodium-glucose co-transporter-2 inhibitors) within the follow-up period of major cardiovascular outcomes trials (CVOT)

State a single effect measure (e.g., relative risk or risk difference): Hazard ratio

1: Is the analysis of effect modification based on comparison within rather than between trials?

[**X**] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Subgroup analysis or meta-regression comparing overall effects of each individual trial. This is typical for aggregate data meta-analysis.	*Subgroup analysis or meta-regression with most information coming from overall effects, but some trials providing within-trial subgroup information*	*Most trials providing within-trial subgroup information; or individual participant data analysis that combines within and between trial information*	*All trials providing within-trial subgroup information or individual participant data; and the analysis separates within from between trial information, e.g., meta-analysis of interactions*

Comment: This network meta-regression incorporated major CVOT to investigate potential cardiovascular benefit of SGL-T2I in patients with type 2 diabetes (T2D).

2: For within-trial comparisons, is the effect modification similar from trial to trial? [ ] Not applicable: no or one within-RCT comparison

[ ] Definitely not similar	[**X**] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Effect modification reported for two or more trials and clearly different directions	*Effect modification not reported for individual trials or too imprecise to tell*	*Effect modification reported for two or more trials, mostly similar in direction, but considerable differences in magnitude*	*Effect modification reported for two or more trials, similar in direction, only some differences in magnitude*

Comment: There has not been any post hoc analysis investigating whether GLP-1RA can reduce cardiovascular death in patients with prior MI

3: For between-trial comparisons, is the number of trials large? [ ] Not applicable: no between RCT comparison

[ ] Very small	[**X**] Rather small or unclear	[ ] Rather large	[ ] Large
1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 19 CVOT of GLP-1RA and SGLT-2I (including 6 post hoc analyses) were included in the network meta-regression. However, the data points are still considered scarce, and the regression model might be over specified.

4: Was the direction of effect modification correctly hypothesized a priori?

[ ] Definitely no	[**X**] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of effect modification*	*Prior protocol available and includes correct specification of direction of effect modification, e.g., based on a biologic rationale*

Comment: No information.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification? (consider irrespective of number of effect modifiers)

[ ] Chance a very likely explanation	[X] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The mean and 95% CI of the coefficient (β) [-0.07 (-0.34, 0.20)] suggest that the correlation – effect modification of prior MI on cardiovascular death in patients receiving SGLT-2I may be negative.

6: Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[X] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of effect modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 effect modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer effect modifiers tested*	*Protocol available and 3 or fewer effect modifiers tested or number considered in analysis*

Comment: Only effect modifications of the co-treatment of GLP-RA/SGLT-2I and prior cardiovascular diseases were assessed on cardiovascular death in patients with T2D, though the protocol is unavailable.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[X] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: A random effect Bayesian model was applied in the network meta-regression.

8: If the effect modifier is a continuous variable, were arbitrary cut points avoided? [X] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior RCT*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment:

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [X] not applicable

	[ ] Yes, probably decrease	[ ] Yes, probably increase	

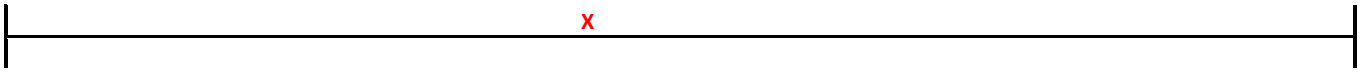
Comment:

10: How would you rate the overall credibility of the proposed effect modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Minimal to no support for effect modification; Use overall effect for each subgroup	Some but insufficient support for effect modification; Use overall effect for each subgroup but note remaining uncertainty	Likely effect modification; Use separate effects for each subgroup but note remaining uncertainty	Very likely effect modification; Use separate effects for each subgroup

Comment: Consistency across studies was unclear. Prior knowledge was unclear. The regression model might become over specified due to scarcity of the data points. The β was generated with less statistical power (because of the over-specification of the model).

Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in a meta-analysis of randomized controlled trials

Version 1.1

Consider the following important instructions informed by common misapplications of ICEMAN in studies using the instrument

- Complete a separate credibility assessment per each effect modifier (e.g., age, comorbidity, drug dose, etc.), outcome (e.g., mortality, stroke, duration of hospital stay), time-point (e.g., 3 months, 6 months), and effect measure (e.g. relative risk, risk difference).
- Do not apply ICEMAN if the interaction p-value is 0.1 or larger, i.e., provides very little statistical support for the existence of an effect modification (ICEMAN is designed to address the possible claim of an effect modification rather than the claim of no effect modification).
- Response options on the left indicate definitely or probably reduced credibility, response options on the right probably or definitely increased credibility
- Completely unclear should be interpreted as probably reduced credibility.
- To ensure transparency, provide a supporting comment under each question that provides a rationale for the rating.
- To ensure transparency, provide a copy of the completed ICEMAN instrument in the supplement of your article.

CREDIBILITY ASSESSMENT

Essential preliminary considerations to define the possible effect modification of interest

State a single candidate effect modifier (e.g., age or comorbidity): Prior heart failure (HF)

Was the effect modifier measured before or at randomization? [**X**] yes, continue [ ] no, stop here and refer to manual for further instructions

State a single outcome and time-point (e.g., mortality at 1 year follow-up): Cardiovascular death in patients with T2D receiving glucagon-like peptide-1 receptor agonists (GLP-1RA) within the follow-up period of major cardiovascular outcomes trials (CVOT)

State a single effect measure (e.g., relative risk or risk difference): Hazard ratio

1: Is the analysis of effect modification based on comparison within rather than between trials?

[**X**] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Subgroup analysis or meta-regression comparing overall effects of each individual trial. This is typical for aggregate data meta-analysis.	*Subgroup analysis or meta-regression with most information coming from overall effects, but some trials providing within-trial subgroup information*	*Most trials providing within-trial subgroup information; or individual participant data analysis that combines within and between trial information*	*All trials providing within-trial subgroup information or individual participant data; and the analysis separates within from between trial information, e.g., meta-analysis of interactions*

Comment: This network meta-regression incorporated major CVOT to investigate potential cardiovascular benefit of GLP-1RA in patients with type 2 diabetes (T2D).

2: For within-trial comparisons, is the effect modification similar from trial to trial? [ ] Not applicable: no or one within-RCT comparison

[ ] Definitely not similar	[**X**] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Effect modification reported for two or more trials and clearly different directions	*Effect modification not reported for individual trials or too imprecise to tell*	*Effect modification reported for two or more trials, mostly similar in direction, but considerable differences in magnitude*	*Effect modification reported for two or more trials, similar in direction, only some differences in magnitude*

Comment: There has not been any post hoc analysis investigating whether GLP-1RA can reduce cardiovascular death in patients with prior HF.

3: For between-trial comparisons, is the number of trials large? [ ] Not applicable: no between RCT comparison

[ ] Very small	[**X**] Rather small or unclear	[ ] Rather large	[ ] Large
1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 16 CVOT of GLP-1RA and SGLT-2I (including 3 post hoc analyses) were included in the network meta-regression. However, the data points are still considered scarce, and the regression model might be over specified.

4: Was the direction of effect modification correctly hypothesized a priori?

[ ] Definitely no	[**X**] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of effect modification*	*Prior protocol available and includes correct specification of direction of effect modification, e.g., based on a biologic rationale*

Comment: No information.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification? (consider irrespective of number of effect modifiers)

[ ] Chance a very likely explanation	[X] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The mean and 95% CI of the coefficient (β) [0.05 (-0.24, 0.39)] suggest that the correlation – effect modification of prior HF on cardiovascular death in patients receiving GLP-1RA may be positive.

6: Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[X] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of effect modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 effect modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer effect modifiers tested*	*Protocol available and 3 or fewer effect modifiers tested or number considered in analysis*

Comment: Only effect modifications of the co-treatment of GLP-RA/SGLT-2I and prior cardiovascular diseases were assessed on cardiovascular death in patients with T2D, though the protocol is unavailable.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[X] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: A random effect Bayesian model was applied in the network meta-regression.

8: If the effect modifier is a continuous variable, were arbitrary cut points avoided? [X] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior RCT*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment:

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [X] not applicable

	[ ] Yes, probably decrease	[ ] Yes, probably increase	

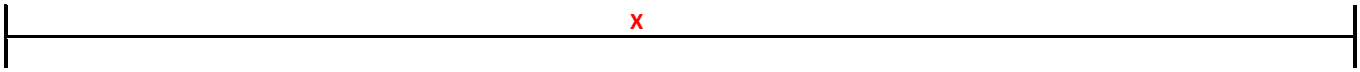
Comment:

10: How would you rate the overall credibility of the proposed effect modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Minimal to no support for effect modification; Use overall effect for each subgroup	Some but insufficient support for effect modification; Use overall effect for each subgroup but note remaining uncertainty	Likely effect modification; Use separate effects for each subgroup but note remaining uncertainty	Very likely effect modification; Use separate effects for each subgroup

Comment: Consistency across studies was unclear. Prior knowledge was unclear. The regression model might become over specified due to scarcity of the data points. The β was generated with less statistical power (because of the over-specification of the model).

Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in a meta-analysis of randomized controlled trials

Version 1.1

Consider the following important instructions informed by common misapplications of ICEMAN in studies using the instrument

- Complete a separate credibility assessment per each effect modifier (e.g., age, comorbidity, drug dose, etc.), outcome (e.g., mortality, stroke, duration of hospital stay), time-point (e.g., 3 months, 6 months), and effect measure (e.g. relative risk, risk difference).
- Do not apply ICEMAN if the interaction p-value is 0.1 or larger, i.e., provides very little statistical support for the existence of an effect modification (ICEMAN is designed to address the possible claim of an effect modification rather than the claim of no effect modification).
- Response options on the left indicate definitely or probably reduced credibility, response options on the right probably or definitely increased credibility
- Completely unclear should be interpreted as probably reduced credibility.
- To ensure transparency, provide a supporting comment under each question that provides a rationale for the rating.
- To ensure transparency, provide a copy of the completed ICEMAN instrument in the supplement of your article.

CREDIBILITY ASSESSMENT

Essential preliminary considerations to define the possible effect modification of interest

State a single candidate effect modifier (e.g., age or comorbidity): Prior heart failure (HF)

Was the effect modifier measured before or at randomization? [**X**] yes, continue [ ] no, stop here and refer to manual for further instructions

State a single outcome and time-point (e.g., mortality at 1 year follow-up): Cardiovascular death in patients with T2D receiving sodium-glucose co-transporter-2 inhibitors (SGLT-2I) within the follow-up period of major cardiovascular outcomes trials (CVOT)

State a single effect measure (e.g., relative risk or risk difference): Hazard ratio

1: Is the analysis of effect modification based on comparison within rather than between trials?

[**X**] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Subgroup analysis or meta-regression comparing overall effects of each individual trial. This is typical for aggregate data meta-analysis.	*Subgroup analysis or meta-regression with most information coming from overall effects, but some trials providing within-trial subgroup information*	*Most trials providing within-trial subgroup information; or individual participant data analysis that combines within and between trial information*	*All trials providing within-trial subgroup information or individual participant data; and the analysis separates within from between trial information, e.g., meta-analysis of interactions*

Comment: This network meta-regression incorporated major CVOT to investigate potential cardiovascular benefit of SGL-2I in patients with type 2 diabetes (T2D).

2: For within-trial comparisons, is the effect modification similar from trial to trial? [ ] Not applicable: no or one within-RCT comparison

[ ] Definitely not similar	[**X**] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Effect modification reported for two or more trials and clearly different directions	*Effect modification not reported for individual trials or too imprecise to tell*	*Effect modification reported for two or more trials, mostly similar in direction, but considerable differences in magnitude*	*Effect modification reported for two or more trials, similar in direction, only some differences in magnitude*

Comment: A meta-analysis of SGLT-2I has indicated that the effect modifications of prior HF on cardiovascular death may be neutral in patients receiving SGLT-2I.

3: For between-trial comparisons, is the number of trials large? [ ] Not applicable: no between RCT comparison

[ ] Very small	[**X**] Rather small or unclear	[ ] Rather large	[ ] Large
1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 16 CVOT of GLP-1RA and SGLT-2I (including 3 post hoc analyses) were included in the network meta-regression. However, the data points are still considered scarce, and the regression model might be over specified.

4: Was the direction of effect modification correctly hypothesized a priori?

[ ] Definitely no	[**X**] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of effect modification*	*Prior protocol available and includes correct specification of direction of effect modification, e.g., based on a biologic rationale*

Comment: No information.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification? (consider irrespective of number of effect modifiers)

[ ] Chance a very likely explanation	[ ] Chance a likely explanation or unclear	[X] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The mean and 95% CI of the coefficient (β) [0.08 (-0.20, 0.37)] suggest that the correlation – effect modification of prior HF on cardiovascular death in patients receiving SGLT-2I may be positive.

6: Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[X] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of effect modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 effect modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer effect modifiers tested*	*Protocol available and 3 or fewer effect modifiers tested or number considered in analysis*

Comment: Only effect modifications of the co-treatment of GLP-RA/SGLT-2I and prior cardiovascular diseases were assessed on cardiovascular death in patients with T2D, though the protocol is unavailable.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[X] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: A random effect Bayesian model was applied in the network meta-regression.

8: If the effect modifier is a continuous variable, were arbitrary cut points avoided? [X] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior RCT*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment:

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [X] not applicable

[ ] Yes, probably decrease	[ ] Yes, probably increase

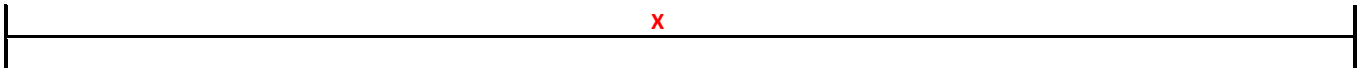
Comment:

10: How would you rate the overall credibility of the proposed effect modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Minimal to no support for effect modification; Use overall effect for each subgroup	Some but insufficient support for effect modification; Use overall effect for each subgroup but note remaining uncertainty	Likely effect modification; Use separate effects for each subgroup but note remaining uncertainty	Very likely effect modification; Use separate effects for each subgroup

Comment: Consistency across studies was unclear. The regression model might become over specified due to scarcity of the data points. The β was generated with less statistical power (because of the over-specification of the model).

Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in a meta-analysis of randomized controlled trials

Version 1.1

Consider the following important instructions informed by common misapplications of ICEMAN in studies using the instrument

- Complete a separate credibility assessment per each effect modifier (e.g., age, comorbidity, drug dose, etc.), outcome (e.g., mortality, stroke, duration of hospital stay), time-point (e.g., 3 months, 6 months), and effect measure (e.g. relative risk, risk difference).
- Do not apply ICEMAN if the interaction p-value is 0.1 or larger, i.e., provides very little statistical support for the existence of an effect modification (ICEMAN is designed to address the possible claim of an effect modification rather than the claim of no effect modification).
- Response options on the left indicate definitely or probably reduced credibility, response options on the right probably or definitely increased credibility
- Completely unclear should be interpreted as probably reduced credibility.
- To ensure transparency, provide a supporting comment under each question that provides a rationale for the rating.
- To ensure transparency, provide a copy of the completed ICEMAN instrument in the supplement of your article.

CREDIBILITY ASSESSMENT

Essential preliminary considerations to define the possible effect modification of interest

State a single candidate effect modifier (e.g., age or comorbidity): Prior heart failure (HF)

Was the effect modifier measured before or at randomization? [**X**] yes, continue [ ] no, stop here and refer to manual for further instructions

State a single outcome and time-point (e.g., mortality at 1 year follow-up): Hospitalization for HF in patients with T2D receiving glucagon-like peptide-1 receptor agonists (GLP-1RA) within the follow-up period of major cardiovascular outcomes trials (CVOT)

State a single effect measure (e.g., relative risk or risk difference): Hazard ratio

1: Is the analysis of effect modification based on comparison within rather than between trials?

[**X**] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Subgroup analysis or meta-regression comparing overall effects of each individual trial. This is typical for aggregate data meta-analysis.	*Subgroup analysis or meta-regression with most information coming from overall effects, but some trials providing within-trial subgroup information*	*Most trials providing within-trial subgroup information; or individual participant data analysis that combines within and between trial information*	*All trials providing within-trial subgroup information or individual participant data; and the analysis separates within from between trial information, e.g., meta-analysis of interactions*

Comment: This network meta-regression incorporated major cardiovascular outcomes trials (CVOT) to investigate potential cardiovascular benefit of GLP-1RA in patients with type 2 diabetes (T2D).

2: For within-trial comparisons, is the effect modification similar from trial to trial? [ ] Not applicable: no or one within-RCT comparison

[ ] Definitely not similar	[**X**] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Effect modification reported for two or more trials and clearly different directions	*Effect modification not reported for individual trials or too imprecise to tell*	*Effect modification reported for two or more trials, mostly similar in direction, but considerable differences in magnitude*	*Effect modification reported for two or more trials, similar in direction, only some differences in magnitude*

Comment: There has not been any post hoc analysis investigating whether GLP-1RA can reduce hospitalization for HF in patients with prior HF.

3: For between-trial comparisons, is the number of trials large? [ ] Not applicable: no between RCT comparison

[ ] Very small	[**X**] Rather small or unclear	[ ] Rather large	[ ] Large
1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 16 CVOT of GLP-1RA and SGLT-2I (including 3 post hoc analyses) were included in the network meta-regression. However, the data points are still considered scarce, and the regression model might be over specified.

4: Was the direction of effect modification correctly hypothesized a priori?

[ ] Definitely no	[**X**] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of effect modification*	*Prior protocol available and includes correct specification of direction of effect modification, e.g., based on a biologic rationale*

Comment: No information.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification? (consider irrespective of number of effect modifiers)

[**X**] Chance a very likely explanation	[ ] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The mean and 95% CI of the coefficient (β) [-0.01 (-0.25, 0.23)] suggest that the correlation – effect modification of prior HF on hospitalization for HF in patients receiving GLP-1RA may be neutral.

6: Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[**X**] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of effect modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 effect modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer effect modifiers tested*	*Protocol available and 3 or fewer effect modifiers tested or number considered in analysis*

Comment: Only effect modifications of the co-treatment of GLP-RA/SGLT-2I and prior cardiovascular diseases were assessed on hospitalization for HF in patients with T2D, though the protocol is unavailable.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[**X**] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: A random effect Bayesian model was applied in the network meta-regression.

8: If the effect modifier is a continuous variable, were arbitrary cut points avoided? [**X**] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior RCT*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment:

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [**X**] not applicable

[ ] Yes, probably decrease	[ ] Yes, probably increase

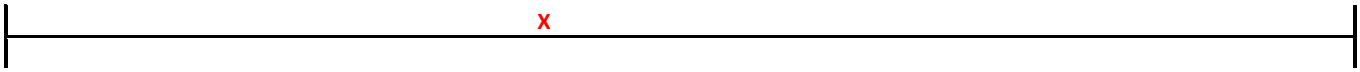
Comment:

10: How would you rate the overall credibility of the proposed effect modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Minimal to no support for effect modification; Use overall effect for each subgroup	Some but insufficient support for effect modification; Use overall effect for each subgroup but note remaining uncertainty	Likely effect modification; Use separate effects for each subgroup but note remaining uncertainty	Very likely effect modification; Use separate effects for each subgroup

Comment: Consistency across studies was unclear. Prior knowledge was unclear. The regression model might become over specified due to scarcity of the data points. The β was generated with less statistical power (because of the over-specification of the model).

Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in a meta-analysis of randomized controlled trials

Version 1.1

Consider the following important instructions informed by common misapplications of ICEMAN in studies using the instrument

- Complete a separate credibility assessment per each effect modifier (e.g., age, comorbidity, drug dose, etc.), outcome (e.g., mortality, stroke, duration of hospital stay), time-point (e.g., 3 months, 6 months), and effect measure (e.g. relative risk, risk difference).
- Do not apply ICEMAN if the interaction p-value is 0.1 or larger, i.e., provides very little statistical support for the existence of an effect modification (ICEMAN is designed to address the possible claim of an effect modification rather than the claim of no effect modification).
- Response options on the left indicate definitely or probably reduced credibility, response options on the right probably or definitely increased credibility
- Completely unclear should be interpreted as probably reduced credibility.
- To ensure transparency, provide a supporting comment under each question that provides a rationale for the rating.
- To ensure transparency, provide a copy of the completed ICEMAN instrument in the supplement of your article.

CREDIBILITY ASSESSMENT

Essential preliminary considerations to define the possible effect modification of interest

State a single candidate effect modifier (e.g., age or comorbidity): Prior heart failure (HF)

Was the effect modifier measured before or at randomization? [**X**] yes, continue [ ] no, stop here and refer to manual for further instructions

State a single outcome and time-point (e.g., mortality at 1 year follow-up): Hospitalization for HF in patients with T2D receiving sodium-glucose co-transporter-2 inhibitors (SGLT-2I) within the follow-up period of major cardiovascular outcomes trials (CVOT)

State a single effect measure (e.g., relative risk or risk difference): Hazard ratio

1: Is the analysis of effect modification based on comparison within rather than between trials?

[**X**] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Subgroup analysis or meta-regression comparing overall effects of each individual trial. This is typical for aggregate data meta-analysis.	*Subgroup analysis or meta-regression with most information coming from overall effects, but some trials providing within-trial subgroup information*	*Most trials providing within-trial subgroup information; or individual participant data analysis that combines within and between trial information*	*All trials providing within-trial subgroup information or individual participant data; and the analysis separates within from between trial information, e.g., meta-analysis of interactions*

Comment: This network meta-regression incorporated major CVOT to investigate potential cardiovascular benefit of SGLT-2I in patients with type 2 diabetes (T2D).

2: For within-trial comparisons, is the effect modification similar from trial to trial? [ ] Not applicable: no or one within-RCT comparison

[ ] Definitely not similar	[**X**] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Effect modification reported for two or more trials and clearly different directions	*Effect modification not reported for individual trials or too imprecise to tell*	*Effect modification reported for two or more trials, mostly similar in direction, but considerable differences in magnitude*	*Effect modification reported for two or more trials, similar in direction, only some differences in magnitude*

Comment: A meta-analysis of SGLT-2I has indicated that the effect modifications of prior HF on hospitalization for HF may be neutral in patients receiving SGLT-2I.

3: For between-trial comparisons, is the number of trials large? [ ] Not applicable: no between RCT comparison

[ ] Very small	[**X**] Rather small or unclear	[ ] Rather large	[ ] Large
1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 16 CVOT of GLP-1RA and SGLT-2I (including 3 post hoc analyses) were included in the network meta-regression. However, the data points are still considered scarce, and the regression model might be over specified.

4: Was the direction of effect modification correctly hypothesized a priori?

[ ] Definitely no	[**X**] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of effect modification*	*Prior protocol available and includes correct specification of direction of effect modification, e.g., based on a biologic rationale*

Comment: No information.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification? (consider irrespective of number of effect modifiers)

[**X**] Chance a very likely explanation	[ ] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The mean and 95% CI of the coefficient (β) [-0.01 (-0.23, 0.22)] suggest that the correlation – effect modification of prior HF on hospitalization for HF in patients receiving SGLT-2I may be neutral.

6: Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[**X**] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of effect modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 effect modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer effect modifiers tested*	*Protocol available and 3 or fewer effect modifiers tested or number considered in analysis*

Comment: Only effect modifications of the co-treatment of GLP-RA/SGLT-2I and prior cardiovascular diseases were assessed on hospitalization for HF in patients with T2D, though the protocol is unavailable.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[**X**] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: A random effect Bayesian model was applied in the network meta-regression.

8: If the effect modifier is a continuous variable, were arbitrary cut points avoided? [**X**] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior RCT*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment:

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [ ] not applicable

[**X**] Yes, probably decrease	[ ] Yes, probably increase

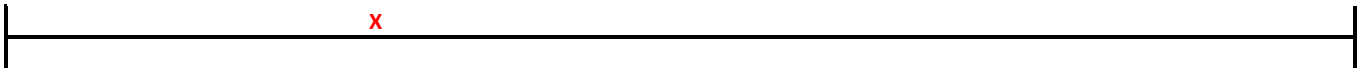
Comment: A meta-analysis of SGLT-2I has indicated that the effect modifications of prior HF on hospitalization for HF may be **neutral** in patients receiving SGLT-2I.

10: How would you rate the overall credibility of the proposed effect modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Minimal to no support for effect modification; Use overall effect for each subgroup	Some but insufficient support for effect modification; Use overall effect for each subgroup but note remaining uncertainty	Likely effect modification; Use separate effects for each subgroup but note remaining uncertainty	Very likely effect modification; Use separate effects for each subgroup

Comment: Consistency across studies was unclear. The regression model might become over specified due to scarcity of the data points. The β was generated with less statistical power (because of the over-specification of the model).

Code for the network meta-analyses and meta-regressions

1. Bayesian network meta-analyses to investigate the cardiovascular benefit of the newer glucose-lowering medications

```
setwd("D:\\GEMTC\\HR network\\ANALYSIS")
library(gemtc)
data <- read.csv("sample.csv", sep=",", header=T)
network <- mtc.network(data.re=data)
model <-mtc.model(network, type ="consistency", n.chain =4, likelihood="binom", link="cloglog",
linearModel="random")
results <- mtc.run(model, n.adapt = 20000, n.iter = 150000, thin = 1)
summary (results)
```

2. Bayesian network meta-regressions to explore the cardiovascular benefit of the combination treatment of GLP-1RA and SGLT-2I

```
setwd("D:\\GEMTC\\HR network\\COMBINATION")
library(gemtc)
data.re <-read.csv("MACE COMBINATION REGRESSION.csv", sep=",", header=T)
combination<-read.csv("MACE COMBINATION.csv", sep=",", header=T)
network <- mtc.network(data.re=data.re, studies= combination)
plot(network)
model<-mtc.model(network, type="regression", regressor=list(coefficient='shared', variable='COM',
control='placebo'), likelihood="binom", link="cloglog")
results <- mtc.run(model, n.adapt = 20000, n.iter = 150000, thin = 1)
summary(results)
```

3. Bayesian network meta-regressions to estimate cardiovascular outcomes of GLP-1RA and SGLT-2I under effect modification of pre-existing cardiovascular diseases

```
setwd("D:\\GEMTC\\HR network\\REGRESSION")
library(gemtc)
data.re <-read.csv("MI MI REGRESSION.csv", sep=",", header=T)
history<-read.csv("MI MI history.csv", sep=",", header=T)
network <- mtc.network(data.re=data.re, studies=history)
plot(network)
model<-mtc.model(network, type="regression", regressor=list(coefficient='unrelated', variable='MI',
control='placebo'), likelihood="binom", link="cloglog")
results <- mtc.run(model, n.adapt = 20000, n.iter = 150000, thin = 1)
summary(results)
```

Cardiorenal outcomes of combination of SGLT-2 inhibitors and GLP-1 receptor agonists in patients with type 2 diabetes: a systematic review and meta-analysis

Citation

John Wilding, JingJing Zhu, XiaoSong Gu. Cardiorenal outcomes of combination of SGLT-2 inhibitors and GLP-1 receptor agonists in patients with type 2 diabetes: a systematic review and meta-analysis. PROSPERO 2022 CRD42022385007 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022385007

Review question

The cardiorenal benefits of SGLT-2 inhibitors or GLP-1 receptor agonists have well been established. Recent clinical trials have demonstrated that combination of SGLT-2 inhibitors and GLP-1 receptor agonists are safe and well tolerated in patients with type 2 diabetes. With respect to efficacy, SGLT-2 inhibitors and GLP-1 receptor agonists can complement each other in reducing HbA1c, systolic blood pressure, total and LDL cholesterol, and weight loss, which suggests that the combination may further improve cardiorenal outcomes in patients with type 2 diabetes, over and above what has been shown for the individual drugs in each class.

The purpose of this systematic review and meta-analysis is to investigate potential cardiorenal benefits of the combination therapy of SGLT-2 inhibitors and GLP-1 receptor agonists in patients with type 2 diabetes, based on subgroup analysis (stratified by the add-on SGLT-2 inhibitors or GLP-1 receptor agonists at baseline or initiated during the trials) to determine whether there are additional effects on specified cardiorenal outcomes in those patients treated with both classes of medication during the trials.

PICOS:

Population: People with type 2 diabetes at high cardiovascular risks;

Interventions: Combination therapy of SGLT-2 inhibitors and GLP-1 receptor agonists;

Comparators: Placebo; patients treated with SGLT-2 inhibitors or GLP-1 receptor agonists;

Outcomes: Major adverse cardiovascular events (MACE: composite of non- fatal myocardial infarction, non-fatal stroke and cardiovascular death); hospitalization for heart failure (HF) or a composite of death from cardiovascular causes or hospitalization for HF; a composite renal outcome as reported in the trials.

Searches

We will search PubMed and EMBASE to identify eligible trials investigating cardiorenal outcomes of SGLT-2 inhibitors or GLP-1 receptor agonists in patients with type 2 diabetes. The search will be filtered to include only randomized controlled trials or suitable post-hoc analysis of these trials completed by 18 DEC 2022. Systematic reviews or meta-analyses will not be included.

Types of study to be included

Randomized clinical trials for cardiovascular or renal outcomes in people with type 2 diabetes.; studies that have included participants without diabetes will not be included.

Condition or domain being studied

Type 2 diabetes, and related cardiovascular and renal outcomes.

Participants/population [1 change]

Inclusion: Adult patients with type 2 diabetes included in eligible clinical trials with primary outcomes of cardiovascular or renal events.

Exclusion: Patients without type 2 diabetes (determined at baseline) in cardiovascular or renal outcome trials of SGLT-2 inhibitors or GLP-1 receptor agonists.

Patients under 18 years of age.

Intervention(s), exposure(s)

The combination therapy of SGLT-2 inhibitors and GLP-1 receptor agonists used together at any time during the trial. The SGLT-2 inhibitors included are dapagliflozin, canagliflozin, empagliflozin, ertugliflozin, sotagliflozin. The GLP-1 receptor agonists included are lixisenatide, semaglutide, liraglutide, exenatide, albiglutide, dulaglutide and efpeglenatide.

Comparator(s)/control

Comparison will be made with treatment outcomes compared with placebo, SGLT-2 inhibitors or GLP-1 receptor agonists. The SGLT-2 inhibitors and GLP-1 receptor agonists are those specified above. Considering DPP-4 inhibitors increase endogenous GLP-1, in randomized clinical trials investigating cardiorenal outcomes of SGLT-2 inhibitors, the cardiorenal outcomes in the combination therapy of SGLT-2 inhibitors and DPP-4 inhibitors (prescribed at baseline or initiated during the trial) will also be used to compare with the those in the intervention – combination therapy of SGLT-2 inhibitors and GLP-1 receptor agonists.

Context

We will only include studies from published clinical trials in people with type 2 diabetes where primary outcomes included MACE, HF or a composite renal outcome.

Main outcome(s) [1 change]

MACE (if sufficient data is available then individual components of the primary outcome will also be assessed);

A composite renal outcome (including 40% or 50% reduction in eGFR, doubling of the serum creatinine level, need for renal replacement therapy, or death from renal causes).

Measures of effect

Hazard ratio (HR) with corresponding 95% confidence interval (CI).

Additional outcome(s) [1 change]

Hospitalization for HF or a composite of death from cardiovascular causes or hospitalization for HF;

A composite renal outcome as indicated in main outcomes above.

Measures of effect

HR with corresponding 95% CI.

Data extraction (selection and coding) [1 change]

Published randomized controlled trials (including their supplementary appendices) and post-hoc analysis of these trials will be consulted for data extraction. Clinical trial investigators will also be approached to retrieve unpublished data.

Data to be retrieved include the name, year of publication, and intervention for each trial. We will then require for each of the subgroups to be assessed (ie those receiving placebo, trial medication alone, or combination with SGLT2i / GLP1 RA), the major baseline characteristics of the patients in each trial (age, sex, history of cardiovascular disease (Y/N), history of heart failure, baseline eGFR, baseline albuminuria), main outcomes of MACE and its individual components; composite renal outcomes as reported in each trial. Additional outcomes of hospitalization for HF and a composite of death from cardiovascular causes or hospitalization for HF will also be ascertained. These will be tabulated in Excel and exported to Stata for analysis.

Data will be extracted by two investigators (JingJing Zhu and XiaoSong Gu). Disagreements will be resolved with consensus by the third investigator (JPHW).

Risk of bias (quality) assessment

The Cochrane Collaboration Risk-of-Bias tool will be used for quality assessment of the eligible trials. Publication bias might be evaluated by funnel plots using Begg's rank test, the Egger's regression test and the trim and fill method.

Strategy for data synthesis

We will conduct a meta-analysis by applying the inverse variance-weighted averages of pooled logarithmic hazard ratio using a random-effects analysis with STATA 16. Sensitivity analysis will be performed to determine whether a single study could affect the aggregate result or not. Heterogeneity will be measured using Higgins I^2 and Cochrane Q statistic. Heterogeneity will be considered as low ($I^2 < 25\%$), moderate (25–50%), high (>50%). A (two-sided) p value of < 0.05 will be considered as statistically significant.

Analysis of subgroups or subsets

In randomized clinical trials investigating cardiorenal outcomes of SGLT-2 inhibitors, subgroups will be stratified based on whether patients received GLP-1 receptor agonists or DDP-4 inhibitors (including those prescribed at baseline or initiated during the trials).

In randomized clinical trials investigating cardiorenal outcomes of GLP-1 receptor agonists, subgroups will be stratified based on whether patients received SGLT-2 inhibitors (including those already being prescribed at baseline or initiated or initiated during the trials).

Contact details for further information

John P. H. Wilding

j.p.h.wilding@liverpool.ac.uk

Organisational affiliation of the review

Department of Cardiovascular and Metabolic Medicine, Institute of Life Course and Medical Sciences, University of Liverpool, United Kingdom

https://www.liverpool.ac.uk/life-course-and-medical-sciences/about/cardiovascular-and-metabolic-medicine/

Review team members and their organisational affiliations [1 change]

Professor John Wilding. University of Liverpool

Dr JingJing Zhu. The Second Affiliated Hospital of Soochow University

Dr XiaoSong Gu. The Second Affiliated Hospital of Soochow University

Type and method of review

Meta-analysis, Systematic review

Anticipated or actual start date [1 change]

01 November 2022

Anticipated completion date

31 March 2023

Funding sources/sponsors

The China Scholarship Council (Grant number: 202006920018)

The Second Affiliated Hospital of Soochow University (Grant number: SDFEYBS1815; XKTJ-HRC2021007)

Soochow University (Grant number: No. GZK1202135)

Grant number(s)

State the funder, grant or award number and the date of award

Conflicts of interest

1. Professor John P. H. Wilding (JPHW) undertakes consultancy for industry contracted via the University of Liverpool (no personal payment) in relation to obesity and type 2 diabetes: AstraZeneca, Boehringer Ingelheim, Eli Lilly, Napp, Novo Nordisk, Mundipharma, Rhythm Pharmaceuticals, Sanofi, and Saniona;

2. JPHW is named grantholder (at University of Liverpool) for research grants for clinical trials from AstraZeneca and Novo Nordisk;

3. JPHW – steering committee member DECLARE TIMI 58; investigator, CANVAS, SUSTAIN 6 trials that are included in this systematic review.

Yes

Language

English

Country

China, England

Stage of review

Review Ongoing

Subject index terms status

Subject indexing assigned by CRD

Subject index terms

Cardiovascular Diseases; Cholesterol, LDL; Diabetes Mellitus, Type 2; Glucagon-Like Peptide-1 Receptor; Glycated Hemoglobin; Heart Failure; Hospitalization; Humans; Infarction; Risk Factors; Sodium-Glucose Transporter 2 Inhibitors

Date of registration in PROSPERO

21 December 2022

Date of first submission

19 December 2022

Stage of review at time of this submission

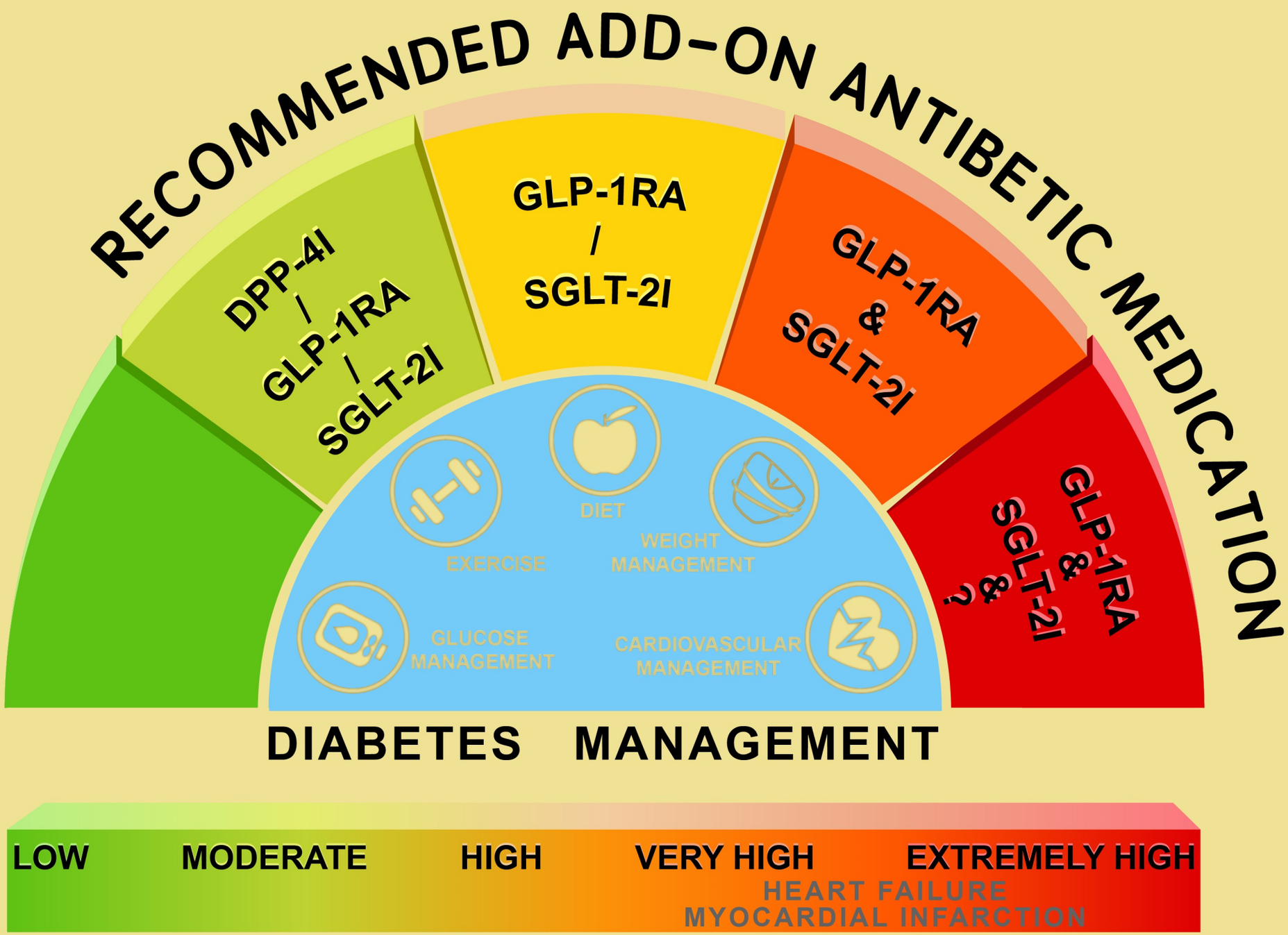
Stage	Started	Completed
Preliminary searches	Yes	Yes
Piloting of the study selection process	Yes	Yes
Formal screening of search results against eligibility criteria	Yes	No
Data extraction	No	No
Risk of bias (quality) assessment	No	No
Data analysis	No	No

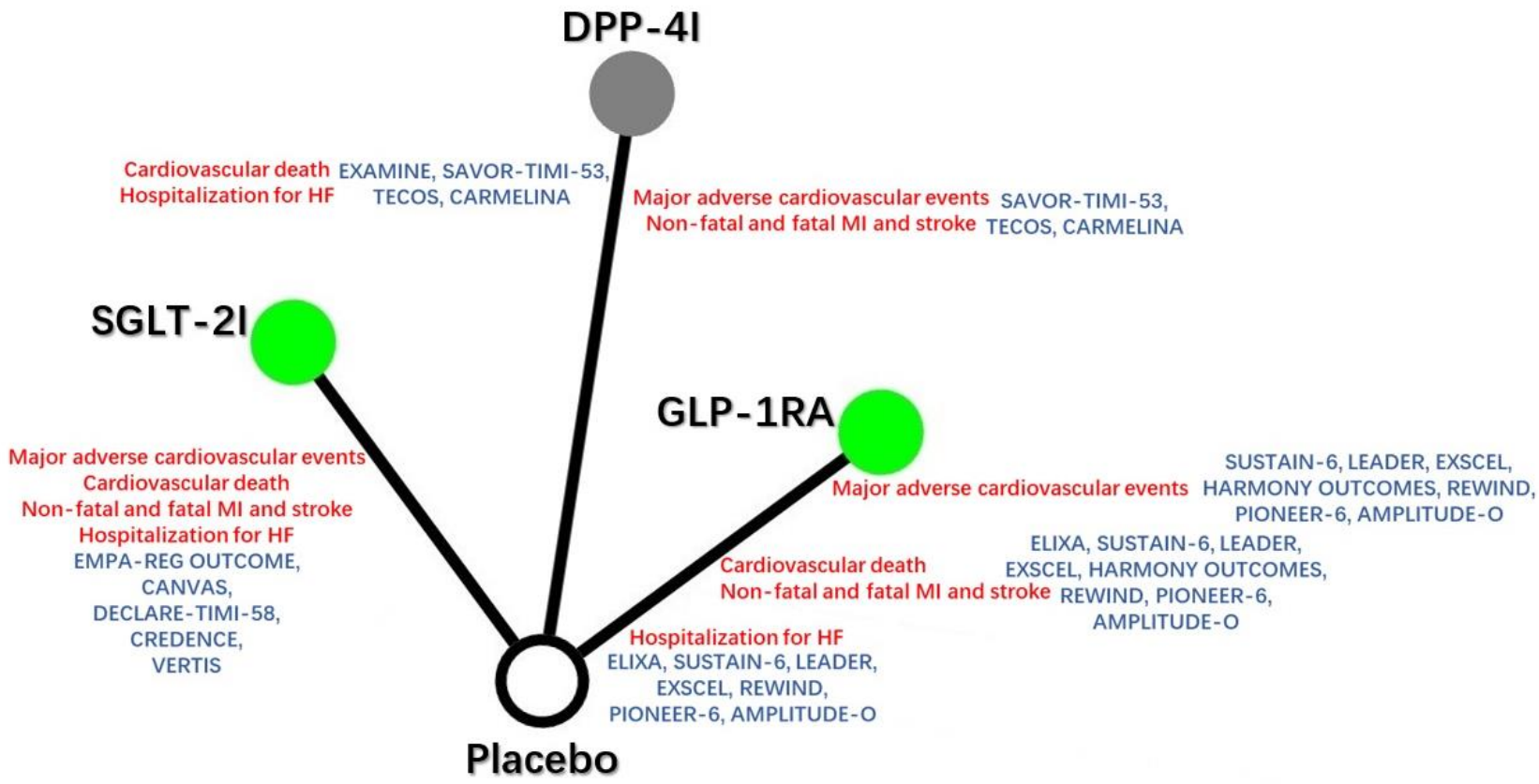
The record owner confirms that the information they have supplied for this submission is accurate and complete and they understand that deliberate provision of inaccurate information or omission of data may be construed as scientific misconduct.

The record owner confirms that they will update the status of the review when it is completed and will add publication details in due course.

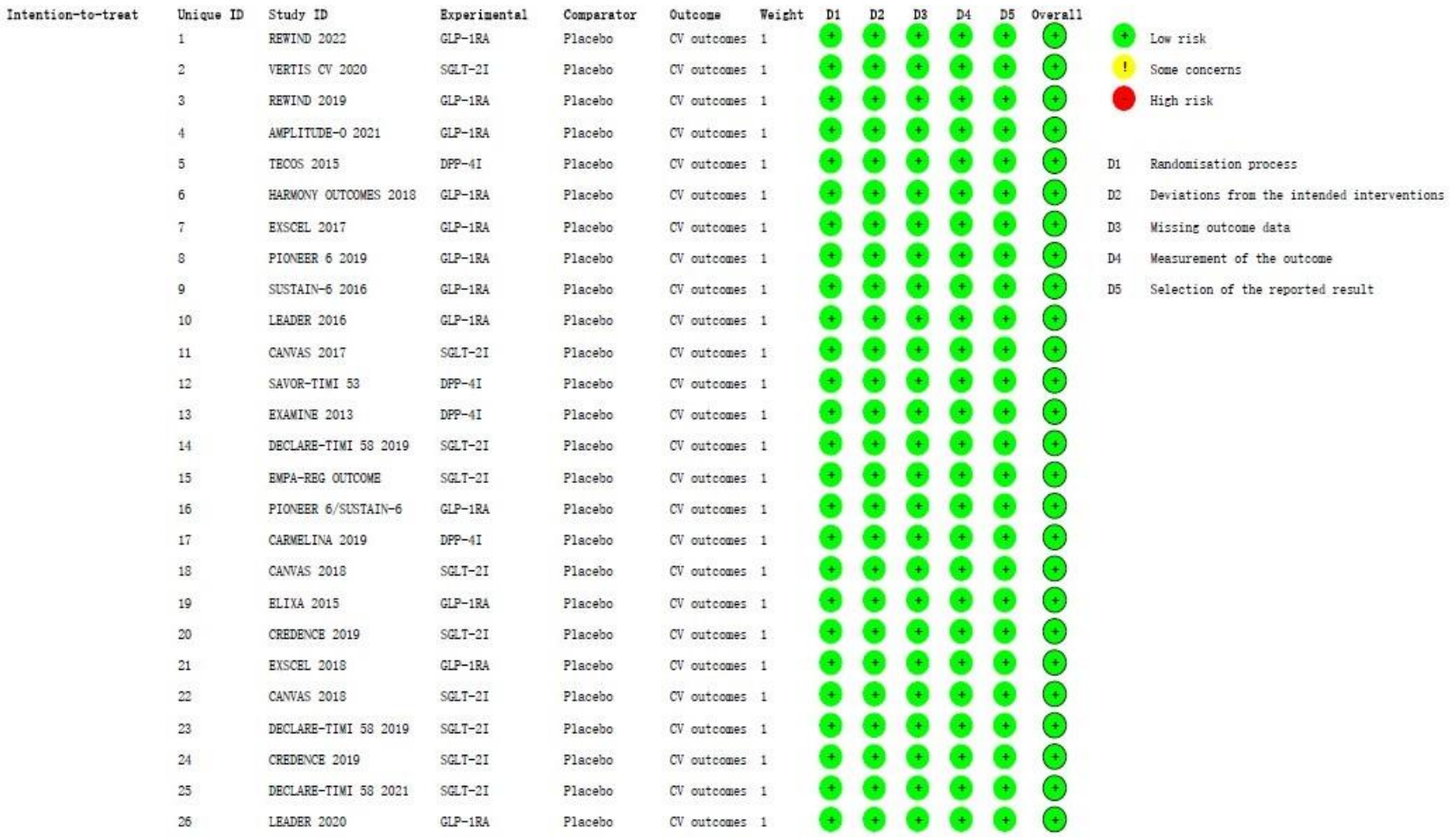
Versions

21 December 2022

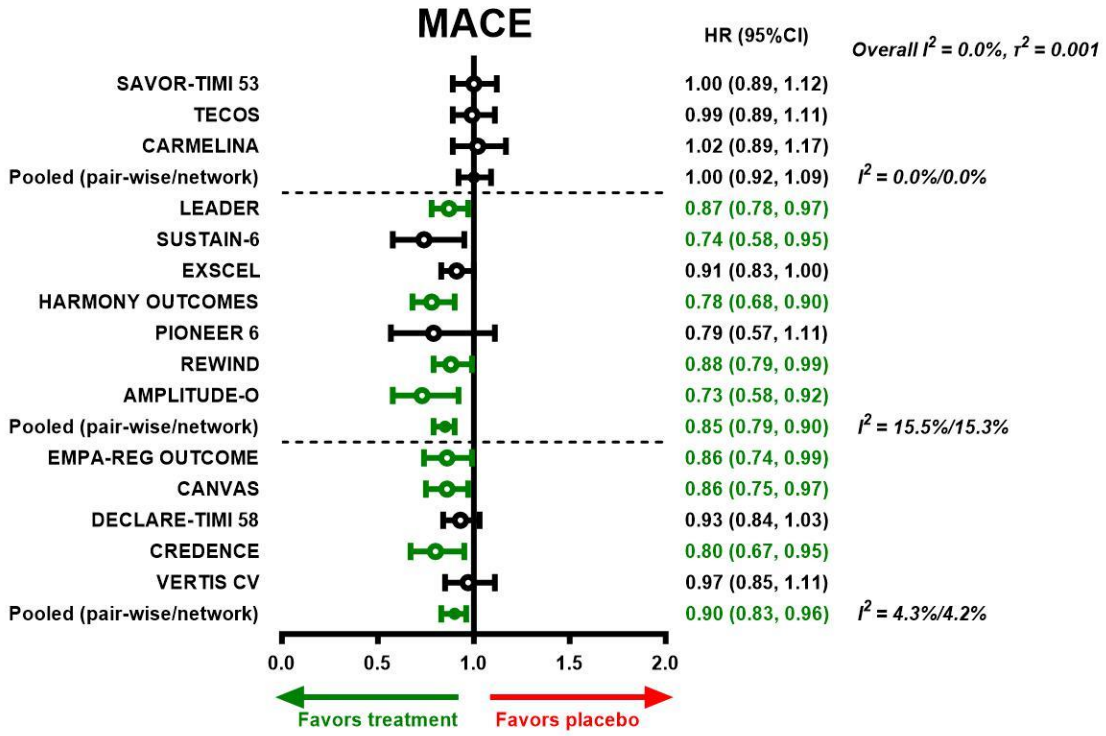

RECOMMENDED ADD-ON ANTIBETIC MEDICATION
DPP-4I / GLP-1RA / SGLT-2I
GLP-1RA / SGLT-2I
GLP-1RA & SGLT-2I
GLP-1RA & SGLT-2I & ?
DIET
EXERCISE
WEIGHT MANAGEMENT
GLUCOSE MANAGEMENT
CARDIOVASCULAR MANAGEMENT
DIABETES MANAGEMENT
LOW
MODERATE
HIGH
VERY HIGH
EXTREMELY HIGH
HEART FAILURE
MYOCARDIAL INFARCTION
CARDIOVASCULAR RISK SPECTRUM FOR TYPE 2 DIABETES



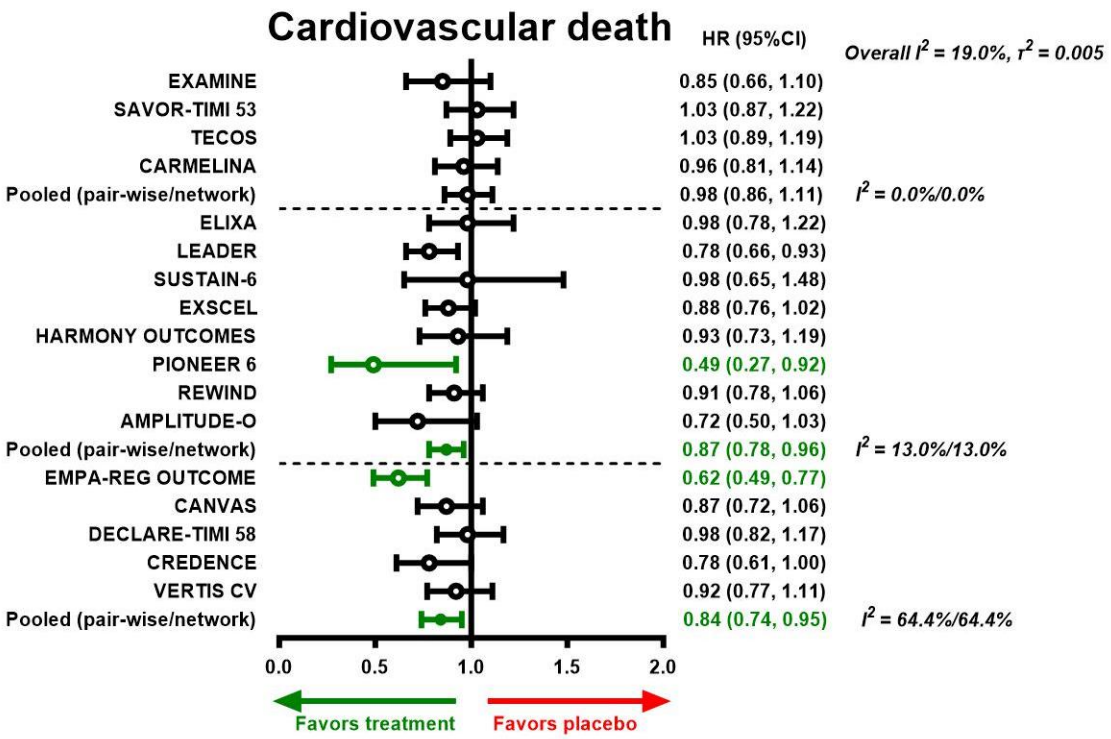
Supplementary Figure 1 The network to investigate the cardiovascular benefit of the newer glucose-lowering medications All the eligible CVOT comparing the cardiovascular outcomes of DPP-4I, GLP-1RA and SGLT-2I were double-blind, randomized placebo-controlled trials. For each treatment, the CVOT included in the network meta-analyses for each cardiovascular outcome are indicated in blue and red, respectively. Treatment with and without potential cardiovascular benefit are indicated in green and grey, respectively.



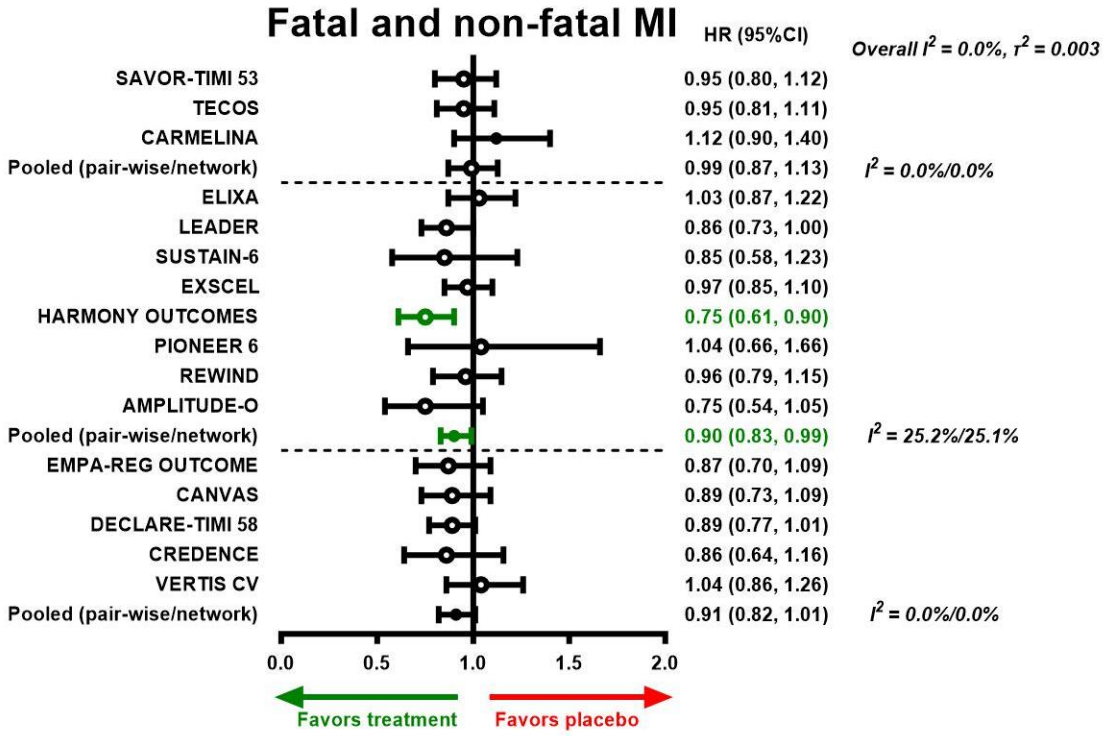
Supplementary Figure 2 Summary of risk of bias Bias of all eligible CVOT (including 17 primary investigations and 9 post hoc analyses (*References*))were assessed as low risk in 5 domains using the Cochrane Collaboration Risk-of-Bias tool.



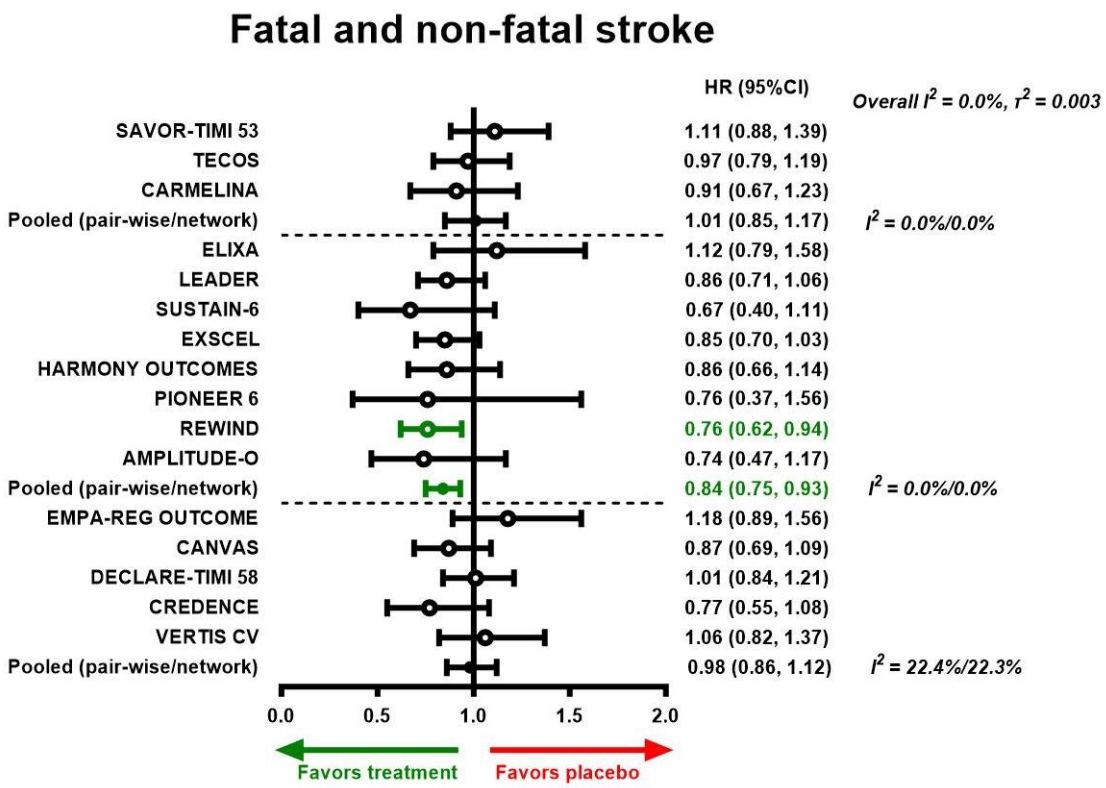
Supplementary Figure 3 Heterogeneity in MACE $HR_{Treatment\ vs.\ Placebo}$ with 95% CI for the observed MACE were extracted from the CVOT and presented in the forest plots. The pooled HR $_{Treatment\ vs.\ Placebo}$ with 95% CI for each treatment were calculated in the network meta-analyses. HR (< 1) favors the treatment whereas HR (> 1) favors the placebo. Significant results are indicated in green and red. The overall between-study heterogeneities were assessed using the I^2 and τ^2 statistic. For each treatment, pairwise and network heterogeneities were also evaluated using the I^2 statistic.



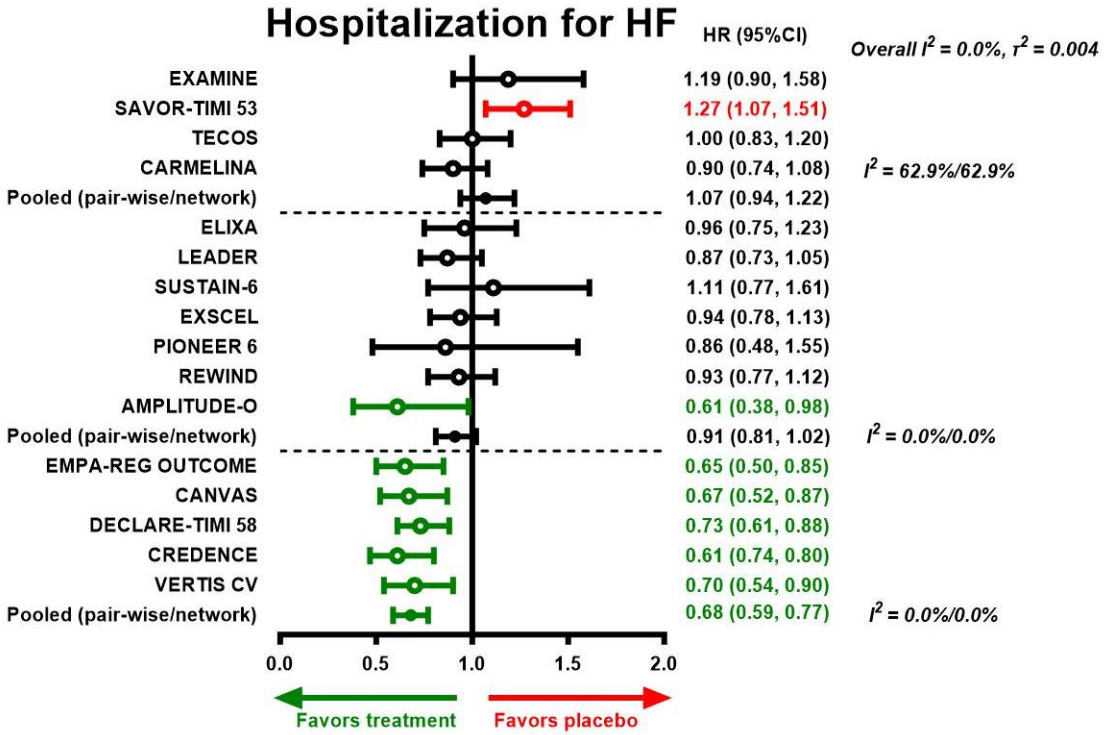
Supplementary Figure 4 Heterogeneity in cardiovascular death $HR_{Treatment\ vs.\ Placebo}$ with 95% CI for the observed cardiovascular death were extracted from the CVOT and presented in the forest plots. The pooled $HR_{Treatment\ vs.\ Placebo}$ with 95% CI for each treatment were calculated in the network meta-analyses. HR (< 1) favors the treatment whereas HR (> 1) favors the placebo. Significant results are indicated in green and red. The overall between-study heterogeneities were assessed using the I^2 and τ^2 statistic. For each treatment, pairwise and network heterogeneities were also evaluated using the I^2 statistic.



Supplementary Figure 5 Heterogeneity in fatal and non-fatal MI $HR_{Treatment\ vs.\ Placebo}$ with 95% CI for the observed fatal and non-fatal MI were extracted from the CVOT and presented in the forest plots. The pooled $HR_{Treatment\ vs.\ Placebo}$ with 95% CI for each treatment were calculated in the network meta-analyses. HR (< 1) favors the treatment whereas HR (> 1) favors the placebo. Significant results are indicated in green and red. The overall between-study heterogeneities were assessed using the I^2 and τ^2 statistic. For each treatment,pairwise and network heterogeneities were also evaluated using the I^2 statistic.



Supplementary Figure 6 Heterogeneity in fatal and non-fatal stroke $HR_{Treatment\ vs.\ Placebo}$ with 95% CI for the observed fatal and non-fatal stroke were extracted from the CVOT and presented in the forest plots. The pooled $HR_{Treatment\ vs.\ Placebo}$ with 95% CI for each treatment were calculated in the network meta-analyses. HR (< 1) favors the treatment whereas HR (> 1) favors the placebo. Significant results are indicated in green and red. The overall between-study heterogeneities were assessed using the I^2 and τ^2 statistic. For each treatment, pairwise and network heterogeneities were also evaluated using the I^2 statistic.



Supplementary Figure 7 Heterogeneity in hospitalization for HF $HR_{Treatment\ vs.\ Placebo}$ with 95% CI for the observed hospitalization for HF were extracted from the CVOT and presented in the forest plots. The pooled HR $_{Treatment\ vs.\ Placebo}$ with 95% CI for each treatment were calculated in the network meta-analyses. HR (< 1) favors the treatment whereas HR (> 1) favors the placebo. Significant results are indicated in green and red. The overall between-study heterogeneities were assessed using the I^2 and τ^2 statistic. For each treatment,pairwise and network heterogeneities were also evaluated using the I^2 statistic.

Supplementary Table 1 The postbaseline co-treatment with GLP-1RA and SGLT-2I are not considered sources of between-study heterogeneity

Cardiovascular outcome	Covariate	I^2	τ^2	β	
				Mean	95% CI
MACE	NA	6%	0.003	NA	NA
	Co-treatment of SGLT-2I/GLP-1RA	4%	0.003	-0.07	(-0.22, 0.08)
Cardiovascular death	NA	6%	0.004	NA	NA
	Co-treatment of SGLT-2I/GLP-1RA	9%	0.005	-0.06	(-0.31, 0.17)
Fatal and non-fatal MI	NA	0%	0.005	NA	NA
	Co-treatment of SGLT-2I/GLP-1RA	0%	0.004	-0.09	(-0.43, 0.11)
Fatal and non-fatal stroke	NA	0%	0.004	NA	NA
	Co-treatment of SGLT-2I/GLP-1RA	0%	0.005	-0.01	(-0.26, 0.24)
Hospitalization for HF	NA	0%	0.004	NA	NA
	Co-treatment of SGLT-2I/GLP-1RA	0%	0.005	-0.13	(-0.42, 0.13)

The covariates of percentages of patients receiving postbaseline co-treatment with SGLT-2I/GLP-1RA were incorporated into the network meta-regressions. The low degrees of variations between I^2 or τ^2 in the network meta-regressions and run-in-parallel meta-analyses (without covariate incorporations) suggest that the co-treatment of SGLT-2I and GLP-1RA are not considered sources of between-study heterogeneity in the CVOT. The estimated means of β with 95% CI indicate that the higher the percentages of patients received the co-treatment, the lower the HR for a particular cardiovascular outcome there might be.

Supplementary Table 2 The baseline cardiovascular co-morbidities are sources of between-study heterogeneity in cardiovascular death

Cardiovascular outcome	Intervention	Covariate	I^2	τ^2	β	
					Mean	95% CI
Fatal and non-fatal MI	GLP-1RA/SGLT-2I vs. Placebo	NA	0%	0.004	NA	NA
	GLP-1RA vs. Placebo	Prior history of MI	0%	0.003	-0.17	(-0.40, 0.05)
	SGLT-2I vs. Placebo				0.06	(-0.11, 0.24)
Cardiovascular death	GLP-1RA/SGLT-2I vs. Placebo	NA	24%	0.011	NA	NA
	GLP-1RA vs. Placebo	Prior history of MI	22%	0.017	-0.06	(-0.43, 0.29)
	SGLT-2I vs. Placebo				-0.07	(-0.34, 0.20)
	GLP-1RA/SGLT-2I vs. Placebo	NA	22%	0.010	NA	NA
	GLP-1RA vs. Placebo	Prior history of HF	18%	0.018	0.05	(-0.24, 0.39)
	SGLT-2I vs. Placebo				0.08	(-0.20, 0.37)
Hospitalization for HF	GLP-1RA/SGLT-2I vs. Placebo	NA	0%	0.002	NA	NA
	GLP-1RA vs. Placebo	Previous history of HF	0%	0.004	-0.01	(-0.25, 0.23)
	SGLT-2I vs. Placebo				-0.01	(-0.23, 0.22)

The covariates of percentages of patients having baseline prior cardiovascular diseases including MI and HF were incorporated in the network meta-regressions. I^2 and τ^2 in cardiovascular death are reduced when covariates of baseline history of MI and HF were incorporated in the network meta-regressions, therefore these preexisting co-morbidities are considered sources of between-study heterogeneity. The estimated means of β with 95% CI indicate that GLP-1RA and SGLT-2I were more effective at prevention of cardiovascular death and hospitalization of HF in trial populations with higher levels of MI and HF, whereas prevention of recurring MI and cardiovascular death became less effective as trial populations of MI and HF increased.