

COMPARING SOC8 TO OTHER CLINICAL GUIDELINES

An in-depth exploration of the SOC-8
methodology

Jon Arcelus,
Co-Chair of SOC-8



26-28 APRIL 2023
KILLARNEY, IRELAND

METHODOLOGY...

2



A pocket watch with a silver chain hangs from the top center. The watch face is white with black numerals and hands. The background consists of several antique books with worn, brown leather covers and gold-tooled decorative patterns.

A LITTLE BIT OF HISTORY

- **LATE 1800S-** LACK LINGUISTIC FRAME TO EXPLAIN TRANS IDENTITIES.
- **BY THE 1940S AND '50S-** EUROPE AND THE U.S: PROFESSIONALS SAW TGD AS PEOPLE WHOSE PSYCHOLOGICAL AND SOMETIMES PHYSICAL MAKE-UP WAS NEITHER PRECISELY MALE NOR FEMALE.
- LAWS CRIMINALIZING SURGEONS WHO REMOVE HEALTHY TISSUE.
- **1960S AND '70S:** EUROPE AND USA: UNIVERSITY GENDER CLINICS WERE ESTABLISHED
- THEY TRY TO ESTABLISH CLINICAL TREATMENT PROTOCOLS.

A LITTLE BIT OF HISTORY

- 1979, THE HARRY BENJAMIN INTERNATIONAL GENDER DYSPHORIA ASSOCIATION WAS FORMED.
 - FIRST “STANDARDS OF CARE [FOR] THE HORMONAL AND SURGICAL SEX REASSIGNMENT OF GENDER DYSPHORIC PERSONS.”
- UPDATED STANDARDS OF CARE WERE PUBLISHED IN 1980, 1981, 1990, 1998, 2001, AND 2011.
- SOC-7 WAS PUBLISHED IN 2012- CONSENSUS BASED (34 PEOPLE)
- SOC ARE WIDELY USED: 232000 RESULTS IN GOOGLE AND OVER 5000 ARTICLES THAT MENTION SOC IN PUBMED
- 5 YEARS AGO, WPATH DECIDED TO INITIATE THE PROCESS OF FOUNDING A NEW SOC-8 REVISION COMMITTEE.

What is a Guideline?

“Guidelines are recommendations intended to assist providers and recipients of health care and other stakeholders to make informed decisions” (WHO)

‘Statements that include recommendations intended to optimise patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options’. (The Institute of Medicine)

Guideline

6

WHY DO WE NEED GUIDELINES?

- TO GUIDE CLINICIANS AND POLICYMAKERS TO IMPROVE CARE, REDUCE VARIATION IN CLINICAL PRACTICE, THEREBY AFFECTING PATIENT SAFETY AND OUTCOMES.

BUT

- RECOMMENDATIONS ARE USED ALONGSIDE PROFESSIONAL JUDGEMENT, DIRECTLY OR WITHIN DECISION AIDS, IN TRAINING AND PRACTICE. THEREFORE, THEY NEED TO BE FLEXIBLE.

*DO CLINICAL
EVIDENCE GUIDELINES
IMPROVE PATIENT
CARE?*

There are no large studies to prove that clinical guidelines improve patient care

CEG present the evidence but do not tell health professional or patients what to do because evidence is only part of making a clinical decision

RECOMMENDED PROCESS FOR GUIDELINES (NICE)



Agree the questions.



Questions define literature searches



A literature search is carried out. This may result in stakeholders providing additional evidence.



A summary of the evidence (known as an evidence review) is prepared.



The impact that the guideline will have on costs is then considered.



The evidence is considered by a committee (Guidelines Development Group) made-up of practitioners, professionals, care providers, commissioners, those who use services and family members or carers.



Question

Develop the question(s)

Evidence

Assess the evidence

Interpret

Interpret the evidence

Recommendations

Make recommendation



LETS LOOK AS TO HOW THE SOC8 WAS DEVELOPED

AIM OF THE SOC8

- **TO PROVIDE CLINICAL GUIDANCE FOR HEALTH PROFESSIONALS**
- **TO ASSIST TRANSGENDER AND GENDER DIVERSE (TGD) PEOPLE WITH SAFE AND EFFECTIVE PATHWAYS TO ACHIEVING LASTING PERSONAL COMFORT WITH THEIR GENDERED SELVES**
- **TO OPTIMIZE THEIR OVERALL PHYSICAL HEALTH, PSYCHOLOGICAL WELL-BEING, AND SELF-FULFILMENT.**



WHO IS THIS FOR?

- ❖ **HEALTH PROFESSIONALS THAT WORK WITH TRANSGENDER AND GENDER DIVERSE PEOPLE**
- ❖ **INCLUDE A DIVERSE ARRAY OF GENDER IDENTITIES AND EXPRESSIONS WHO HAVE DIFFERING NEEDS FOR GENDER AFFIRMING CARE ACROSS THEIR LIFESPAN**
- ❖ **INCLUDE OFTEN NEGLECTED AND/OR MARGINALIZED TGD GROUPS, INCLUDING NONBINARY PEOPLE, EUNUCHS, INDIVIDUALS WITH INTERSEX CONDITIONS, AND THOSE WHO MAY DECIDE TO DETRANSITION.**



13

ADAPTATION OF SOC TO DIFFERENT SETTINGS

- ❖ In some countries the rights of gender minorities are increasingly being recognized and gender clinics are being established
- ❖ Other countries lack such facilities and care may be more fragmented and under-resourced.
- ❖ Different models of care are being pioneered, including efforts to decentralize gender affirming care within primary care settings and telehealth services.
- ❖ Regardless of the care delivery, the principles of gender affirming care as outlined in the SOC should be adapted to local sociocultural, political and medical contexts.



Select the members:

Job descriptions

Application

Selection

Guidelines Committee

Nomination of co-chairs

- Eli Coleman (chair)
- Jon Arcelus & Asa Radix (co-chairs)

Chapter leads

- 50 Applications received
- 18 countries
- 24 Leads chosen

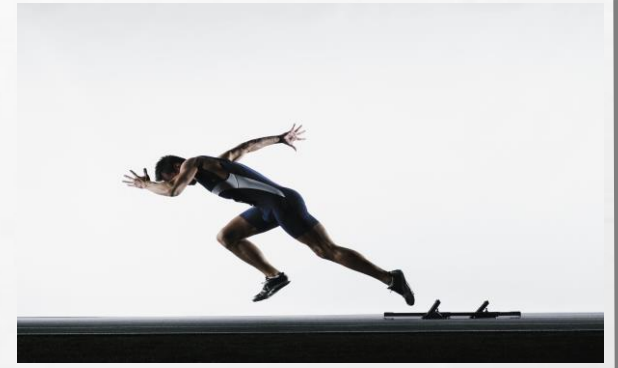
Working groups/stakeholders

- 164 Applications received
- 18 countries
- Team formation

Chapter Members



METHODOLOGY



**Review of SOC7
& Questions**

**Assess
Evidence**

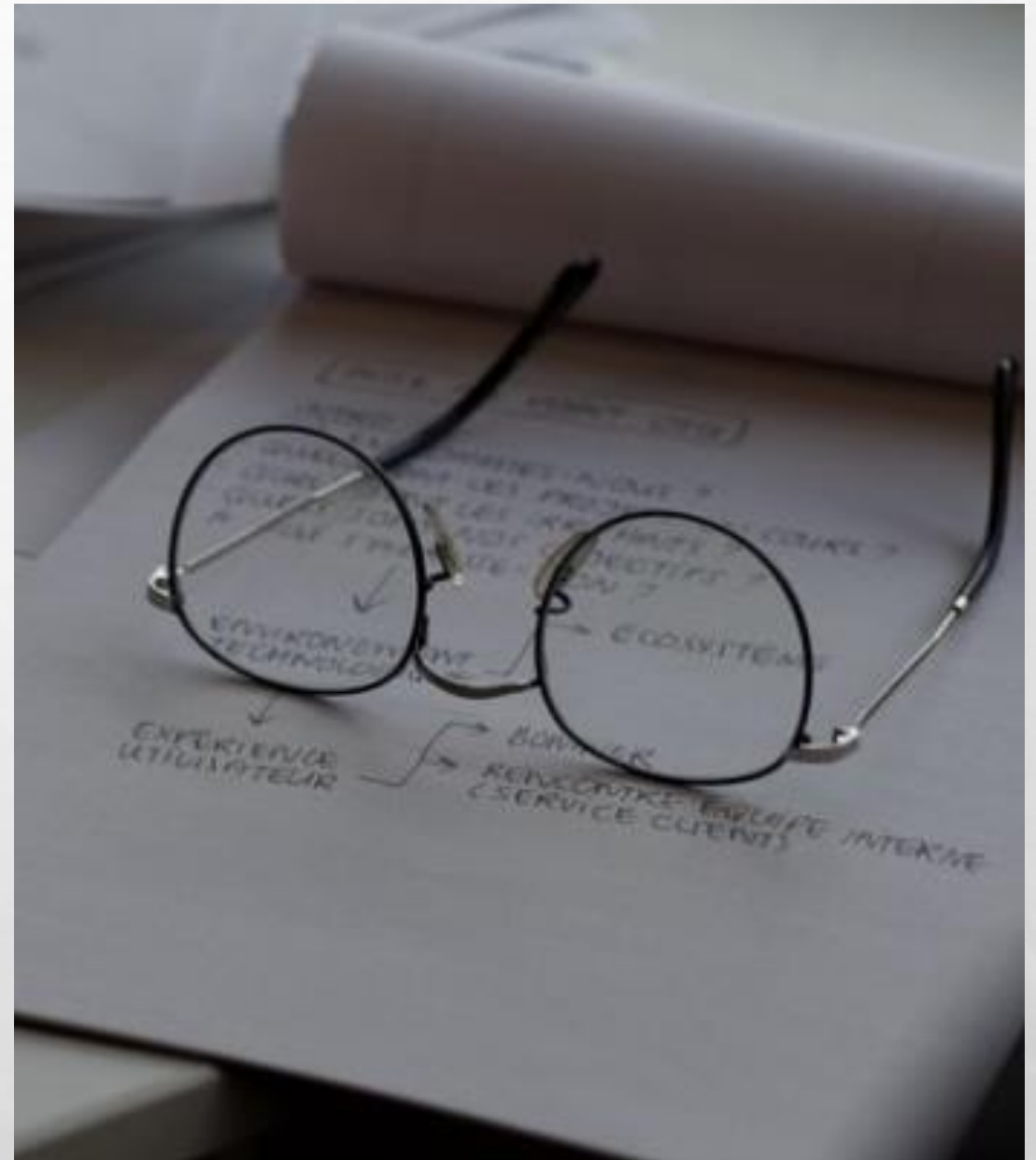
**Develop
Statements**

Delphi

GRADE BASED

REVIEW THE SOC7 AND DEFINE THE CHAPTERS

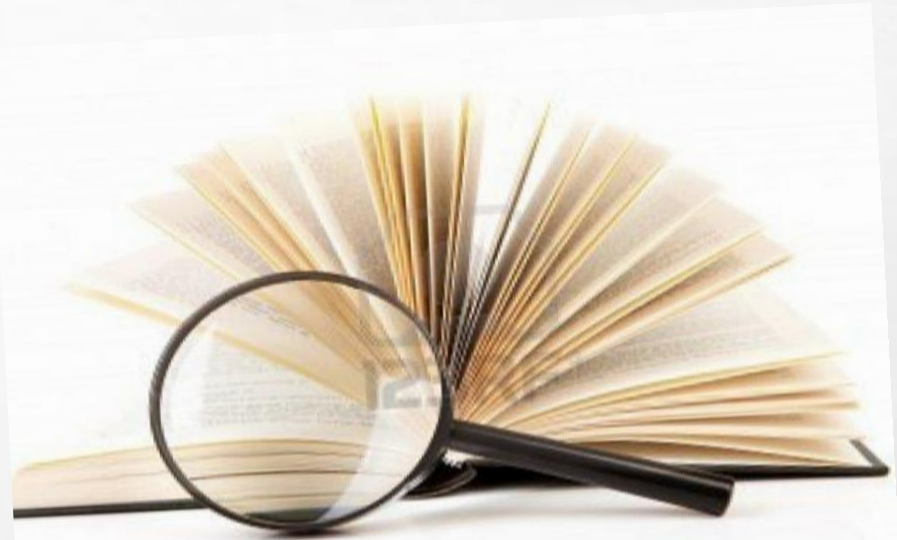
- ❖ REVIEW SOC-7 CHAPTERS
- ❖ REVIEW NEW LITERATURE
- ❖ IDENTIFY NEW AREAS
- ❖ DEFINE THE CHAPTERS



ASSES THE EVIDENCE

SELECT LITERATURE REVIEW INDEPENDENT TEAM:

JOHN HOPKINS UNIVERSITY



Literature Review



RECOMMENDED PROCESS FOR GUIDELINE DEVELOPMENT (NICE)



Agree the questions.



Questions define literature searches



A literature search is carried out. This may result in stakeholders providing additional evidence.



A summary of the evidence (known as an evidence review) is prepared.



The impact that the guideline will have on costs is then considered.



The evidence is considered by a committee (Guidelines Development Group) made-up of practitioners, professionals, care providers, commissioners, those who use services and family members or carers.

RECOMMENDED PROCESS FOR GUIDELINE DEVELOPMENT (NICE)



Agree the questions.



Questions define literature searches



A literature search is carried out. This may result in stakeholders providing additional evidence.



A summary of the evidence (known as an evidence review) is prepared.



The impact that the guideline will have on costs is then considered.



The evidence is considered by a committee (Guidelines Development Group) made-up of practitioners, professionals, care providers, commissioners, those who use services and family members or carers.

DEVELOP THE QUESTIONS:

1- Developed PICO questions for systematic literature reviews

Kq1. For transgender women, what are the safety and efficacy of androgen-lowering medications compared to spironolactone vs cyproterone vs GnRH agonists in terms of surrogate outcomes, clinical outcomes, and harms?

2- Number of PICO questions:

- Adolescents- 2
- Assessment- 4
- Hormone- 13
- Primary care- 5
- Reproduction- 3
- Surgery- 5
- Voice- 8.

PICO

The PICO Principle assists you in organizing and focusing your question into a searchable query.

P

Participants / Population
Who are the relevant patients?

I

Intervention / Indication
What is the management strategy, diagnostic test or exposure you are researching?

C

Comparator / Control
Is there a control or alternative management strategy, test or exposure?

O

Outcome
What are the patient-relevant consequences?

PICO USED FOR MANY CHAPTERS



RECOMMENDED PROCESS FOR GUIDELINE DEVELOPMENT (NICE)



Agree the questions.



Questions define literature searches



A literature search is carried out. This may result in stakeholders providing additional evidence.



A summary of the evidence (known as an evidence review) is prepared.



The impact that the guideline will have on costs is then considered.



The evidence is considered by a committee (Guidelines Development Group) made-up of practitioners, professionals, care providers, commissioners, those who use services and family members or carers.

RECOMMENDED PROCESS FOR GUIDELINE DEVELOPMENT (NICE)



Agree the questions.



Questions
define literature searches



A literature search is
carried out. This may
result in stakeholders
providing additional
evidence.



A summary of the
evidence (known as
an evidence review) is
prepared.



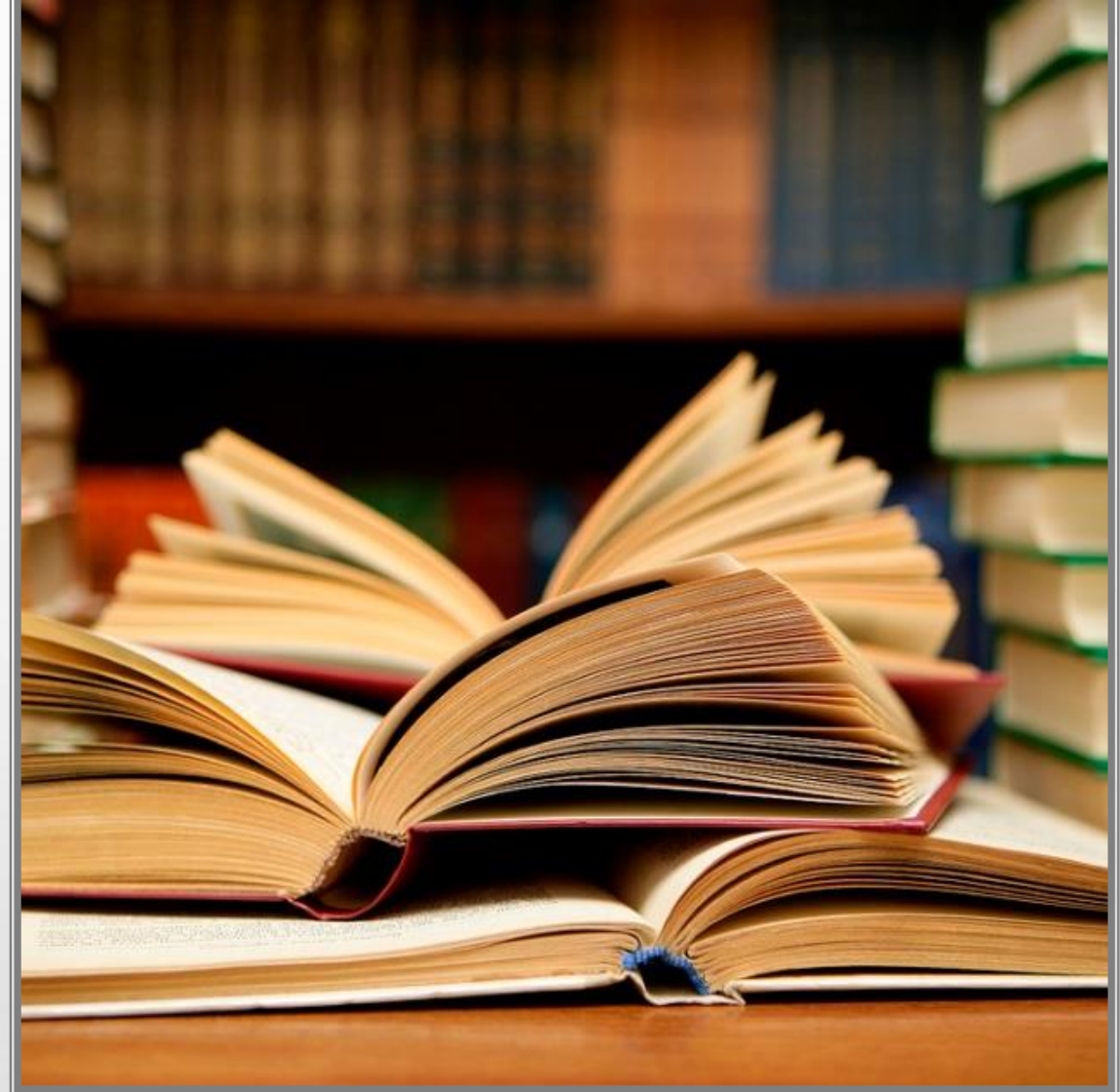
The impact that the
guideline will have
on costs is then
considered.



The evidence
is considered by a
committee (Guidelines
Development Group)
made-up of practitioners,
professionals, care
providers, commissioners,
those who use services
and family members or
carers.

SYSTEMATIC LITERATURE REVIEW

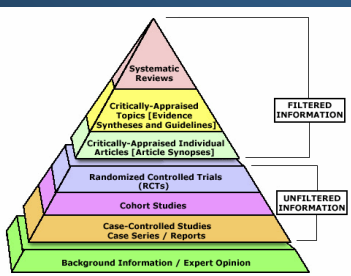
- 37 Independent systematic literature reviews by JHU.
- Every review summarised
- Every study from the SR was assessed for Bias
- Based on this, answers to the PICO questions were provided
- Hundreds of papers reviewed
- Several background reviews by each chapter
- Use of 1800 references



RATE OF EVIDENCE FROM SYSTEMATIC REVIEWS

Evidence could be rate as high, moderate, low and very low

RCTs provide high-certainty estimates, whereas observational studies low certainty



RISK OF BIAS ASSESSED FOR EVERY QUESTION

Author, year	Study design	Confounding	Selection of participants into study	Classification of interventions	Deviations from intended assignments	Missing data	Measurement of outcomes	Selection of reported results	Overall bias
Gava, 2016 ¹	Controlled before-after	Low	Low	Low	NI	NI	Low	Low	Low
Polderman, 1995 ²	Prospective cohort	Moderate	Moderate	Low	Low	Low	Low	Low	Moderate
Cunha, 2018 ³	Retrospective cohort	Moderate	Moderate	Low	NI	Low	Low	Serious	Serious

- RISK OF BIAS OF STUDIES THAT COMPARE THE EFFECTS OF ANDROGEN-LOWERING MEDICATIONS TO SPIRONOLACTONE, CYPROTERONE, OR GONADOTROPIN-RELEASING AGONISTS IN TERMS OF SURROGATE OUTCOMES, CLINICAL OUTCOMES, OR HARMS IN TRANSGENDER WOMEN**

MAIN REASONS FOR BIASES

- **LOW NUMBER IN THE STUDIES**
- **OBSERVATION STUDIES**
- **LOW NUMBER OF FOLLOW UPS**
- **AS MOST STUDIES WERE CLINICAL STUDIES- DEVELOPED WITH LITTLE OR NO FUNDING....**

**BUT THIS IS NOT
VERY DIFFERENT
TO OTHER
CLINICAL AREAS:
E.G. EATING
DISORDERS –
NICE-**

- PICO question: does any psychological intervention produce benefits to people with eating disorders
- Look at over 30 studies- including 16 RCT
- 29 possible outcomes
- Evidence- 7 very low, 22 low
- Do we stop psychological treatments because of this?
- **Recommendations:** for adults with anorexia nervosa consider CBT-ED

NICE
National Institute for
Health and Care Excellence

RECOMMENDED PROCESS FOR GUIDELINE DEVELOPMENT (NICE)



Agree the questions.



Questions define literature searches



A literature search is carried out. This may result in stakeholders providing additional evidence.



A summary of the evidence (known as an evidence review) is prepared.



The impact that the guideline will have on costs is then considered.



The evidence is considered by a committee (Guidelines Development Group) made-up of practitioners, professionals, care providers, commissioners, those who use services and family members or carers.

RECOMMENDED PROCESS FOR GUIDELINE DEVELOPMENT (NICE)



Agree the questions.



Questions define literature searches



A literature search is carried out. This may result in stakeholders providing additional evidence.



A summary of the evidence (known as an evidence review) is prepared.



The impact that the guideline will have on costs is then considered.



The evidence is considered by a committee (Guidelines Development Group) made-up of practitioners, professionals, care providers, commissioners, those who use services and family members or carers.

DEVELOP RECOMMENDATION STATEMENTS OF RECOMMENDATION BASED ON EVIDENCE*

Statements need to be
actionable and
measurable

We are recommending to
professionals (to do
something) no to society

- Eg : People with a chest infection should have antibiotics (NO)
- We recommend that professionals should prescribe antibiotics to people presenting with a chest infection (Yes)

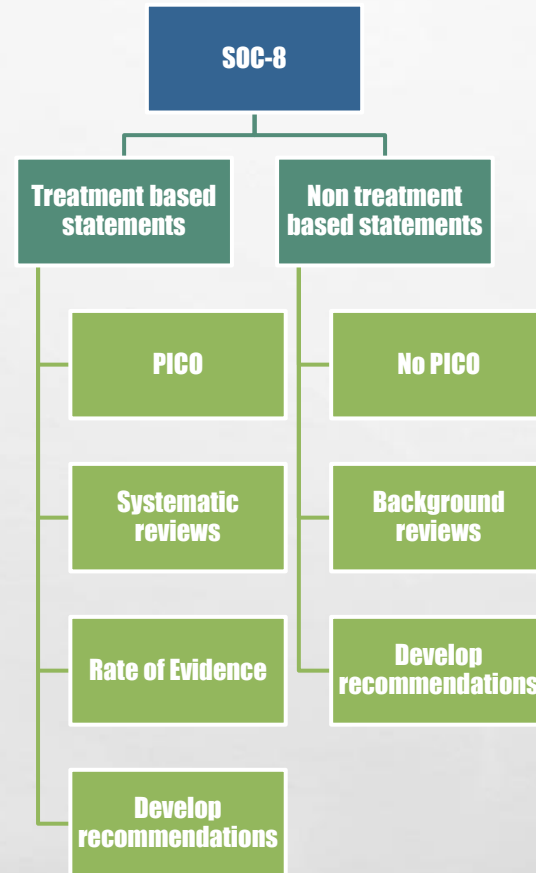
The shorter the statement
and the clearer the better

Statements need to be
recommending something
specific to trans health,
which is different to a
principle of Care.

- Eg. We recommend that clinicians are empathic when working with patients (Principle of care)
- We advise that clinicians should assess the capacity of consent for gender affirming genital surgery prior to recommending this surgery for trans people (statement of recommendation)

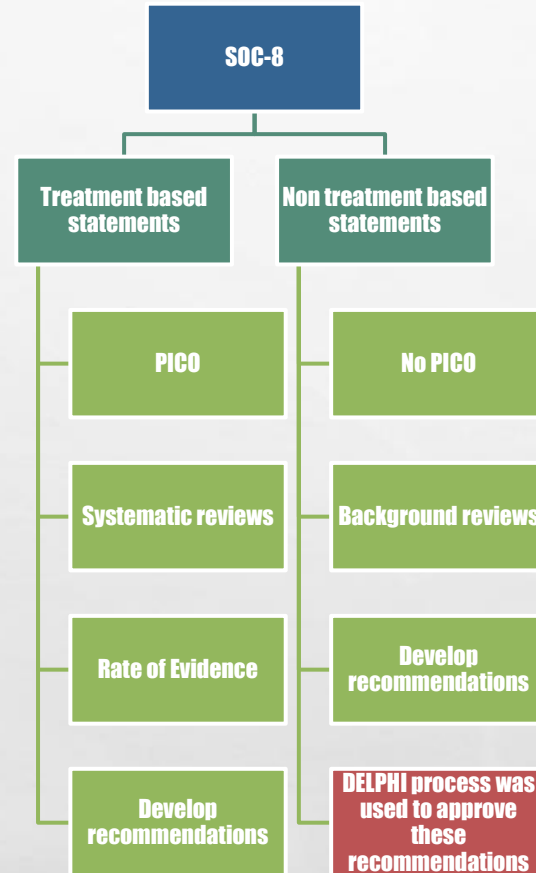
PROCESS OF DEVELOPING RECOMMENDATIONS

RECOMMENDATION:



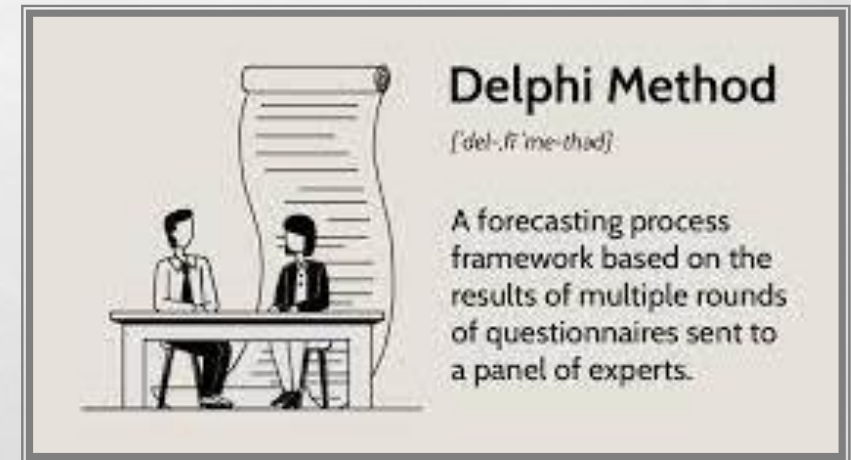
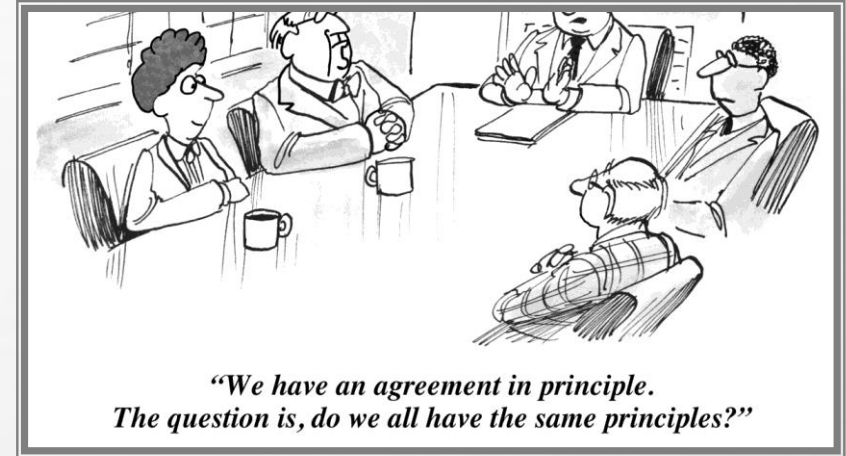
PROCESS OF DEVELOPING RECOMMENDATIONS

RECOMMENDATION:



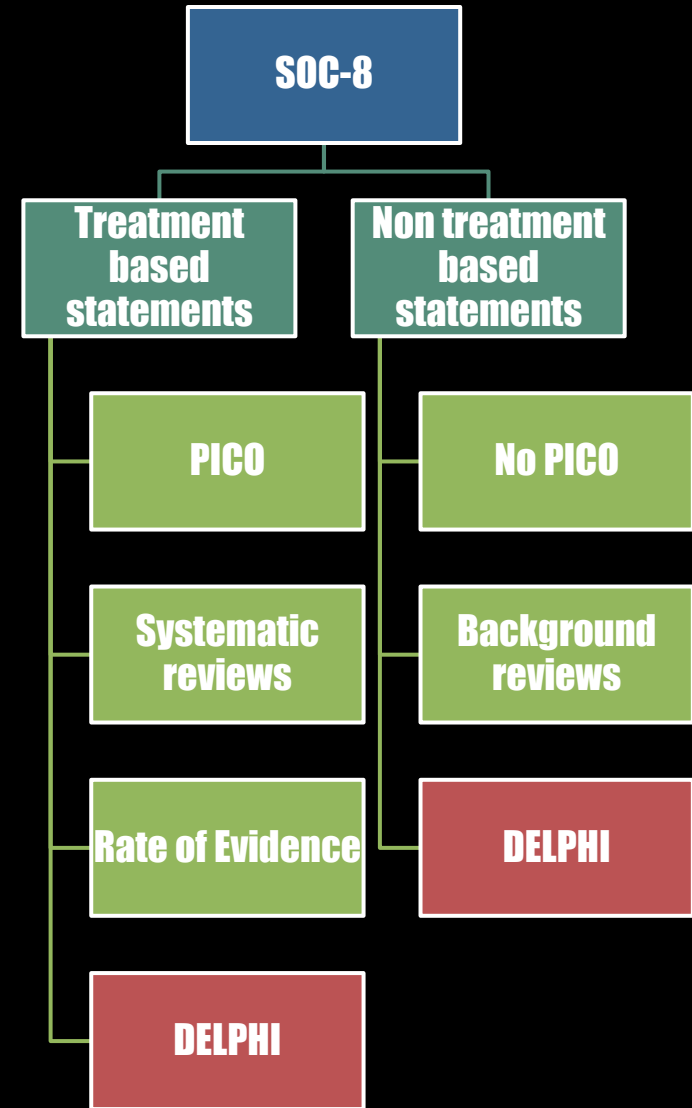
DELPHI

- THE DELPHI METHOD IS A BROADLY ACCEPTED STRATEGY FOR DEVELOPING CONSENSUS RECOMMENDATIONS BASED ON OBJECTIVE EXPERT OPINION
- IT IS INTENDED TO PROVIDE GUIDANCE IN AREAS WHERE LIMITED EVIDENCE-BASED LITERATURE IS AVAILABLE (RAHAGHI ET AL, 2020)



**IN ORDER TO STRENGTH THE
PROCESS WE ALSO INTRODUCED
DELPHI FOR EVERY STATEMENT**

37



VOTING THE STATEMENTS VIA DELPHI

❖ VOTE FOR STATEMENTS VIA DELPHI

❖ 128 PEOPLE

❖ 75% ACCEPTANCE

❖ 3 ROUNDS

❖ ONCE STATEMENT APPROVED, THEY WERE RATED AS STRONG OR WEAK RECOMMENDATION

VOTE



DECIDING WHETHER A RECOMMENDATION WAS STRONG OR WEAK: GRADE



GRADING THE RECOMMENDATION:⁴⁰ GRADE

- *Grading of recommendations, assessment, development and evaluation (GRADE):* developed to address the evidence that is selected and appraised during guidelines development.
- Most widely adopted tool worldwide.
- Strong evidence is usually linked to a strong recommendation
- Expert evidence is of very low quality of evidence by GRADE assessment



GRADE

GRADE CANNOT BE IMPLEMENTED MECHANICALLY – THERE IS BY NECESSITY A CONSIDERABLE AMOUNT OF SUBJECTIVITY IN EACH DECISION.

1. Risk of bias

- **Bias occurs when the results of a study do not represent the truth because of inherent limitations in design or conduct of a study.**

2. Imprecision

- **The GRADE approach to rating imprecision focuses on the 95% confidence interval around the best estimate of the *absolute effect***

3. Inconsistency

- **Certainty in a body of evidence is highest when there are several studies that show consistent effects.**

4. Indirectness

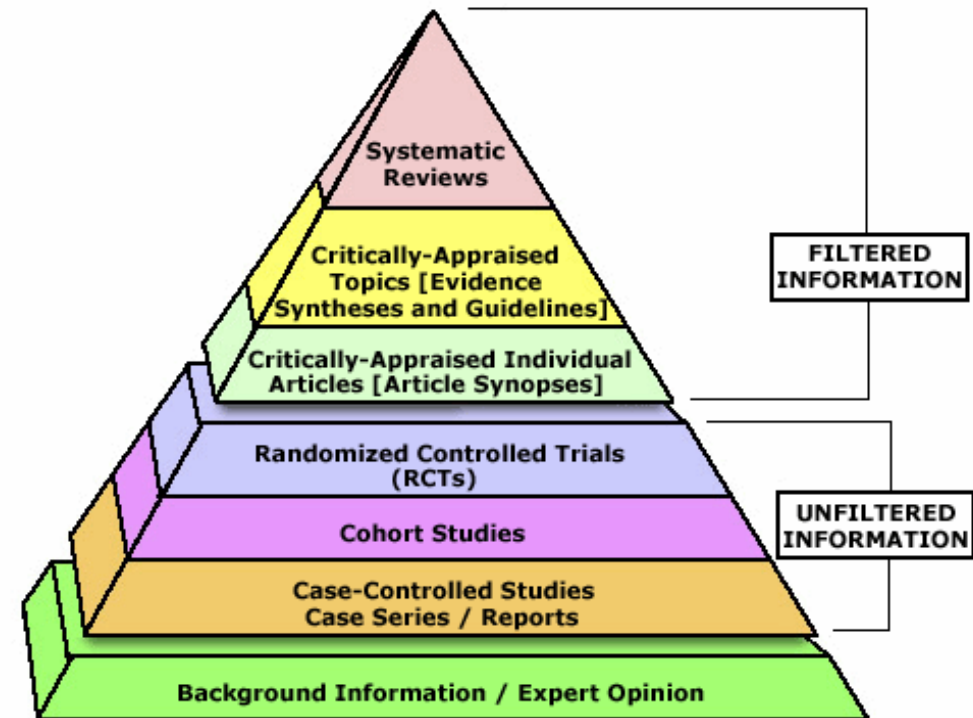
- **Evidence is most certain when studies directly compare the interventions of interest in the population of interest, and report the outcome(s) critical for decision-making.**

5. Publication bias

- **Publication bias is more common with observational data and when most of the published studies are funded by industry.**

**EVIDENCE WAS LOW
OR NON EXISTING
FOR MANY
QUESTIONS-HENCE
WE ALSO USED
EXPERT EVIDENCE**

42



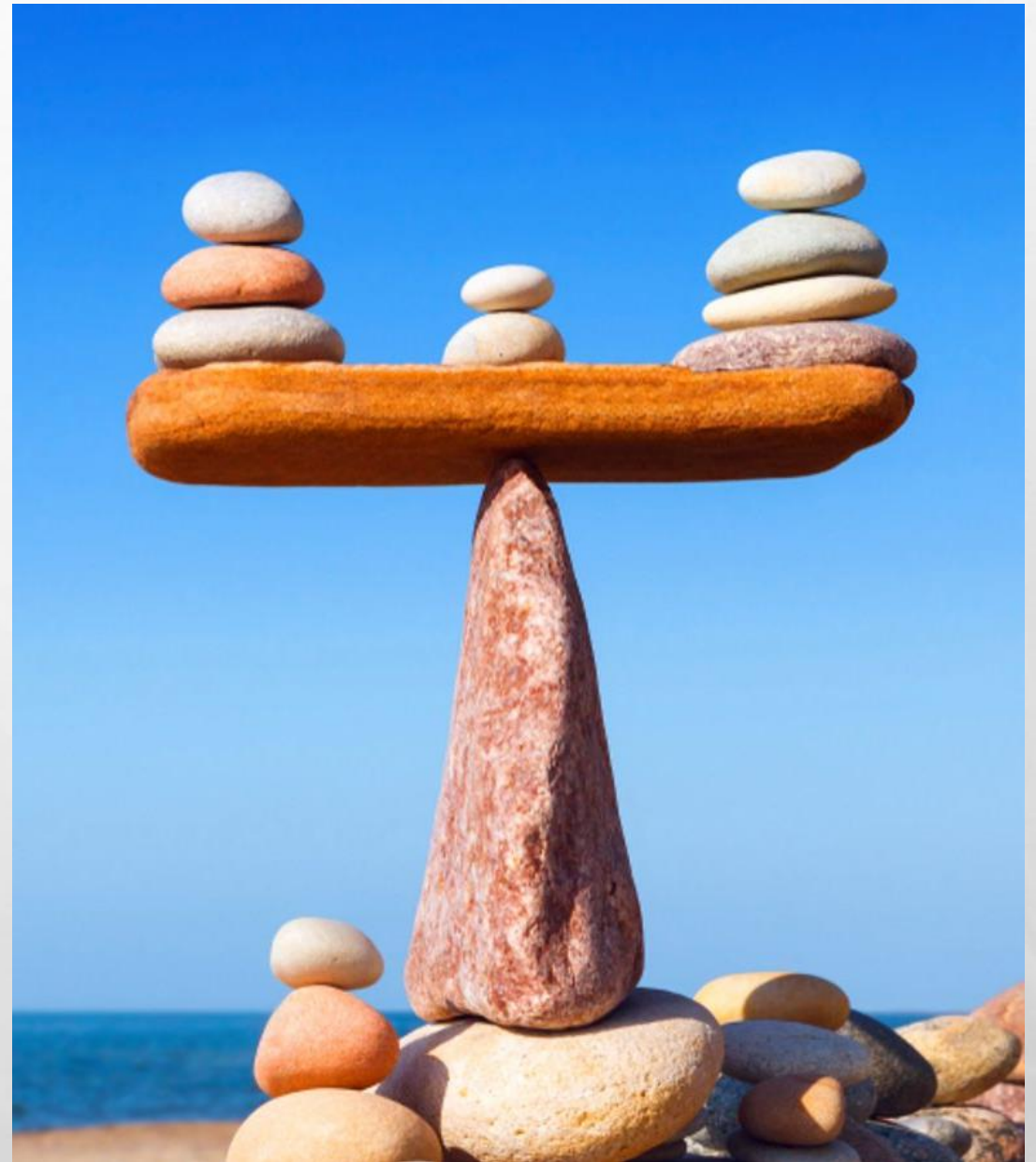
We followed the essence/spirit of EBM, i.e. to follow the best available evidence. So, leading to evidence hierarchy pyramid, when a PICO has only case report or expert opinion as evidence, then that's what you cited as best available evidence (Schünemann et al, 2019)

A photograph of several hands of different skin tones giving thumbs up against a white background. The hands are arranged in a slightly curved line across the upper half of the image.

IT IS GOOD PRACTICE AS GUIDELINE DEVELOPERS THAT YOU STILL FORM A RECOMMENDATION ON THE BASIS OF EXPERT OPINIONS (Neumann & Schünemann (2020)).

GRADING THE RECOMMENDATION BASED ON EXPERT OPINION

- Even if the evidence is low or non-existent the recommendation can still be graded as strong
- Expert evidence is of very low quality of evidence by GRADE assessment, but under some circumstances, low quality of evidence can still support strong recommendations (WHO, 2014)

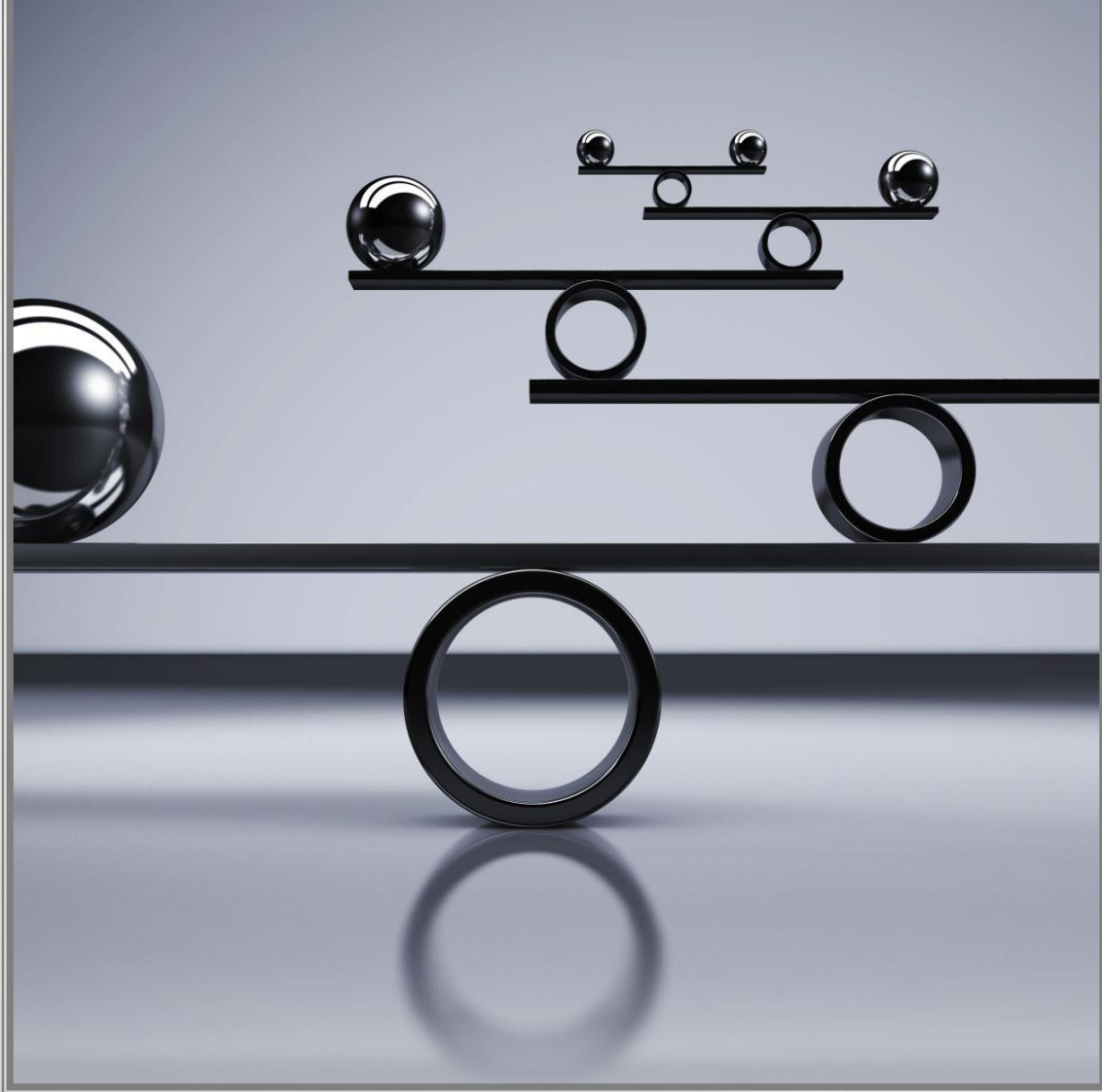


GRADING

When guideline development groups are confident that the desirable consequences (benefits) of an intervention outweigh its undesirable consequences (risks or harms), they will likely issue a strong recommendation (WHO handbook of guideline development, 2014)

Strong recommendations (“we recommend”)

Weak recommendations (“we suggest”)



THE STRENGTH OF RECOMMENDATION CONSIDERS FOUR DOMAINS

- **THE BALANCE OF POTENTIAL BENEFITS AND HARMS**
- **CONFIDENCE IN THAT BALANCE OR QUALITY OF EVIDENCE**
- **VALUES AND PREFERENCES OF PROVIDERS AND PATIENTS**
- **RESOURCE USE AND FEASIBILITY**

NICE GUIDELINES: METHODS



Agree the questions.



Questions define literature searches



A literature search is carried out. This may result in stakeholders providing additional evidence.



A summary of the evidence (known as an evidence review) is prepared.



The impact that the guideline will have on costs is then considered.



The evidence is considered by a committee (Guidelines Development Group) made-up of practitioners, professionals, care providers, commissioners, those who use services and family members or carers.

WRITING SUPPORTIVE TEXT USING LITERATURE



CHECKING REFERENCES, GETTING EXTERNAL FEEDBACK



INTERNATIONAL ORGANIZATIONS PROVIDING FEEDBACK





PUBLIC FEEDBACK PERIOD 8 WEEKS: OVER 2000 COMMENTS

FINAL STAGE

- ❖ COLLECT THE COMMENTS
- ❖ MAKE CHANGES AS APPROPRIATE
- ❖ STATEMENTS WERE NOT MODIFIED UNLESS THEY AFFECT HEALTH CARE OR CONFLICTING WITH OTHER STATEMENTS
- ❖ LAST REVIEW OF THE STRENGTH OF THE RECOMMENDATIONS
- ❖ REVIEW BY EDITOR
- ❖ PUBLICATION





LET'S LOOK AT THE SOC-8 AND COMPARE...



LET'S EXAMINE SOME GUIDELINES RELATED TO TRANSGENDER HEALTH



Consolidated guidelines on HIV prevention, diagnosis, treatment and care for key populations. 2016 update

- WHO (2016)
- Global
- For National HIV programme managers and other decision-makers
- 2 Systematic Reviews
- GRADE,
- Guidelines Development Groups
- 79 independent external peer reviewers

Guidance on oral pre-exposure prophylaxis for serodiscordant couples, men and transgender women who have sex with men at high risk of HIV.

- WHO (2012)
- Global
- For Countries/member states
- 4 Systematic Reviews
- GRADE
- Guidelines development groups
- Independent external review group

Prevention and treatment of HIV and other sexually transmitted infections among men who have sex with men and transgender people. Recommendations for a public health approach

- WHO (2011)
- Global
- For National public health officials and managers of HIV/AIDS and STI programmes, and NGOs
- 13 Systematic Reviews
- GRADE,
- Guidelines Development Group
- Independent external review

ESSM Position Statement
'Assessment and hormonal
management in adolescent and
adult trans people, with attention for
sexual function and satisfaction'

- T'Sjoen et al., 2020
- European Society for Sexual Medicine
- Europe
- 7 authors
- European clinicians working in transgender health, sexologists and other healthcare professionals
- No Systematic Reviews.
- Leading experts' consensus opinion (No GDG).
- No independent external review

Endocrine treatment of gender-dysphoric/ gender-incongruent persons: an Endocrine Society clinical practice guideline

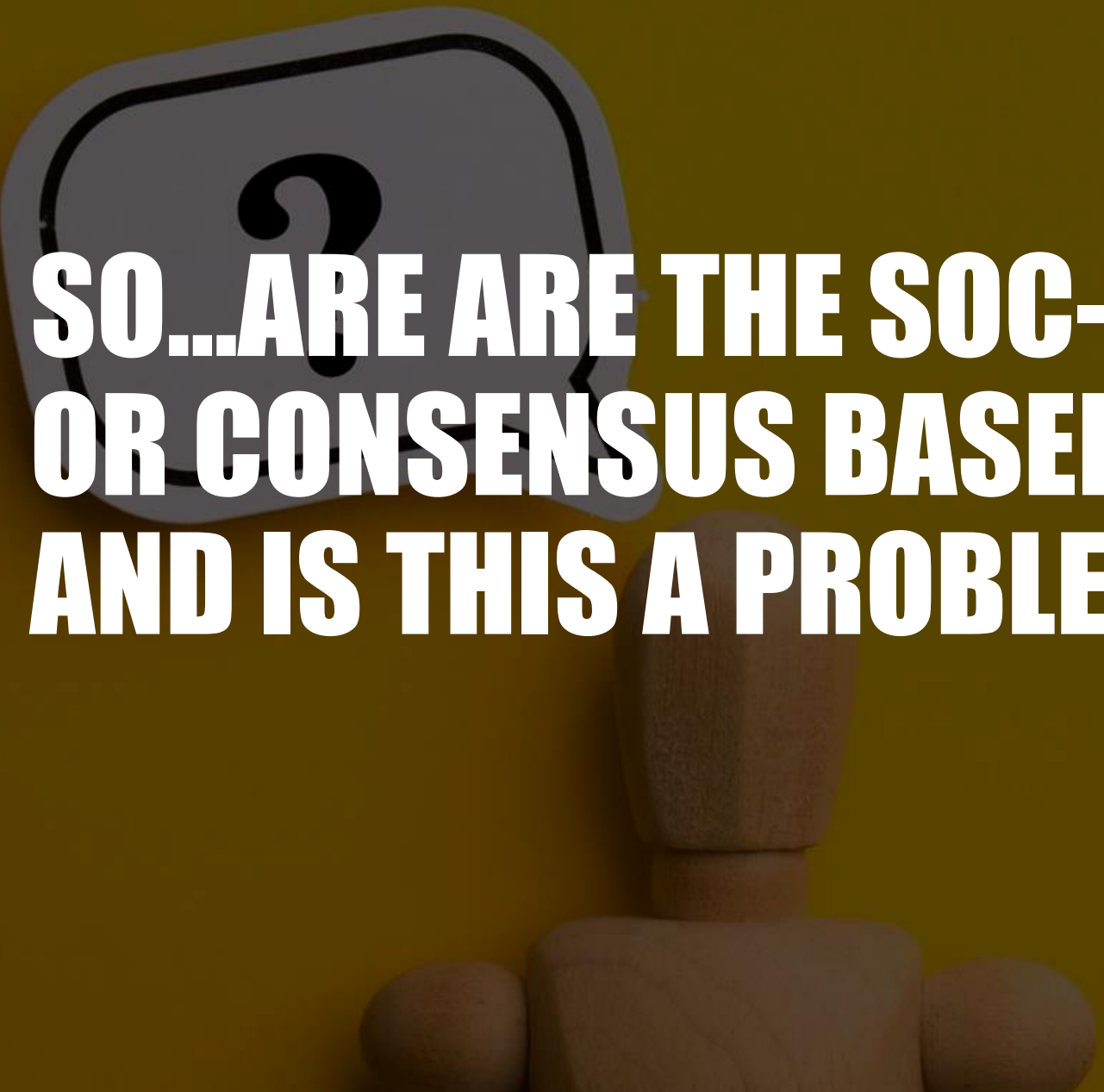
- Hembree et al., 2017
- Endocrine Society
- Global
- 10 authors
- Endocrinologists, mental health professionals and physicians
- 2 Systematic Reviews
- GRADE
- Rest expert consensus.
- No independent external review

SOC-7

- Coleman et al., 2012
- WPATH
- Global
- 34 authors
- For health professionals.
- Work groups submit manuscripts based on prior literature reviews
- No explicit links of recommendations to evidence,
- Expert consensus.
- No independent external review

SOC-8

- Coleman et al., 2022
- WPATH
- Global
- 128 authors
- For health professionals
- Guidelines development groups
- 40 PICO questions
- 37 Independent SR
- Several background reviews with independent monitoring
- Expert consensus based on DELPHI
- GRADE based rates of recommendations
- External feedback

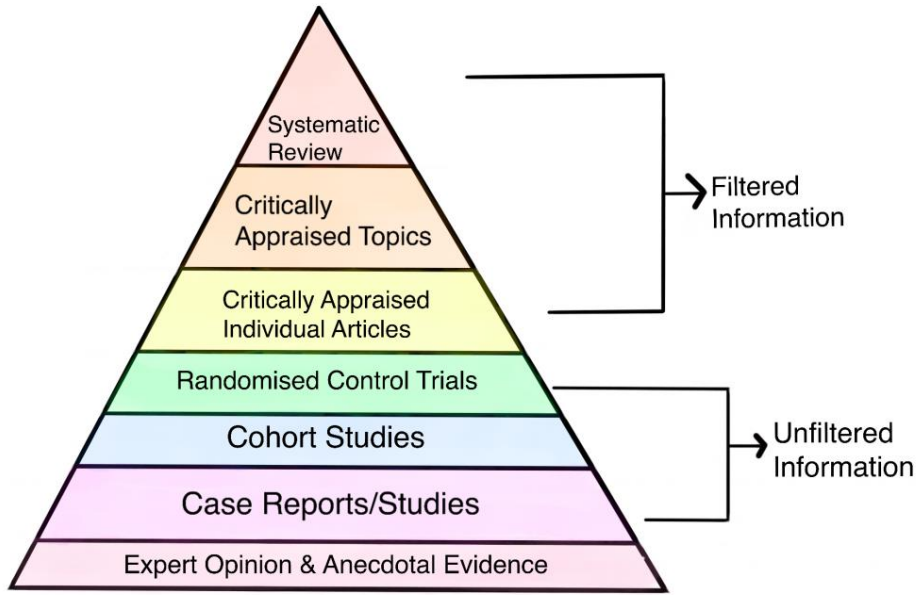
A wooden figure is shown from the chest up, holding a grey question mark sign. The background is a solid olive green color. The text is overlaid on the sign and the figure's torso.

**SO...ARE ARE THE SOC-8 EVIDENCE
OR CONSENSUS BASED?
AND IS THIS A PROBLEM?**



EVIDENCE VS CONSENSUS

- a. Historically, many organizations categorized their guidelines as *evidence-based* or *consensus-based*
- b. When the evidence is of high quality, some guideline panels consider that the evidence speaks for itself and the process is evidence-based.
- c. However, when the evidence is only of low quality or very low quality, some guideline panels label their process as consensus-based.



WHICH ONE IS BEST?

- a. Evidence base is about using the best available evidence to direct clinical decision-making and practice guidelines
- b. Thus, all guidelines should be evidence-based, even when the evidence is of very low quality.
- c. Any guideline development process should include a **systematic review of the literature and rigorous assessment of the quality of the evidence.**
- d. **But the evidence requires interpretation**
- e. On occasion, the evidence is so compelling that answers to such questions are obvious and beyond dispute.
- f. Far more often, the answers are less obvious and require evaluation and, in the context of guidelines, a series of consensus decisions (DELPHI).

Every clinical recommendation, whether from high quality or very low-quality evidence, require consensus from the panel

Even if the evidence is low (expert evidence) it can become a strong recommendation based on the benefit vs the harm review

Making a distinction between evidence-based or consensus-based guidelines is both misguided and misleading

SO...IS THE SOC-8 METHODOLOGY ANY GOOD?

- It follows most of the steps recommended for the development of clinical guidelines
- It has its limitations
 - As the recommendations are not only about treatment
 - There are complex research ethical dilemma to undertake research in the area

LIMITATIONS AND STRENGTHS

Some limitations

1. Selection bias of people as part of the SOC8 group - had to be WPATH members
2. Having specialists making recommendations about their own speciality can add a bias
3. Grading low quality studies and experts opinions

Strengths

- We had people selected based on their applications
- We had stakeholders included
- We had independent literature reviews for systematic literature review
- We had PICO questions
- All PICO questions went through the same Delphi process regardless of the quality of research evidence identified.
- PICO questions went through a rigorous process to identify literature, screen and extract data, even though some PICO questions ended up with little or no evidence
- We had a Delphi process
- We assess the strength of the evidence using GRADE
- We had independent review checkers
- We collected feedback from external sources

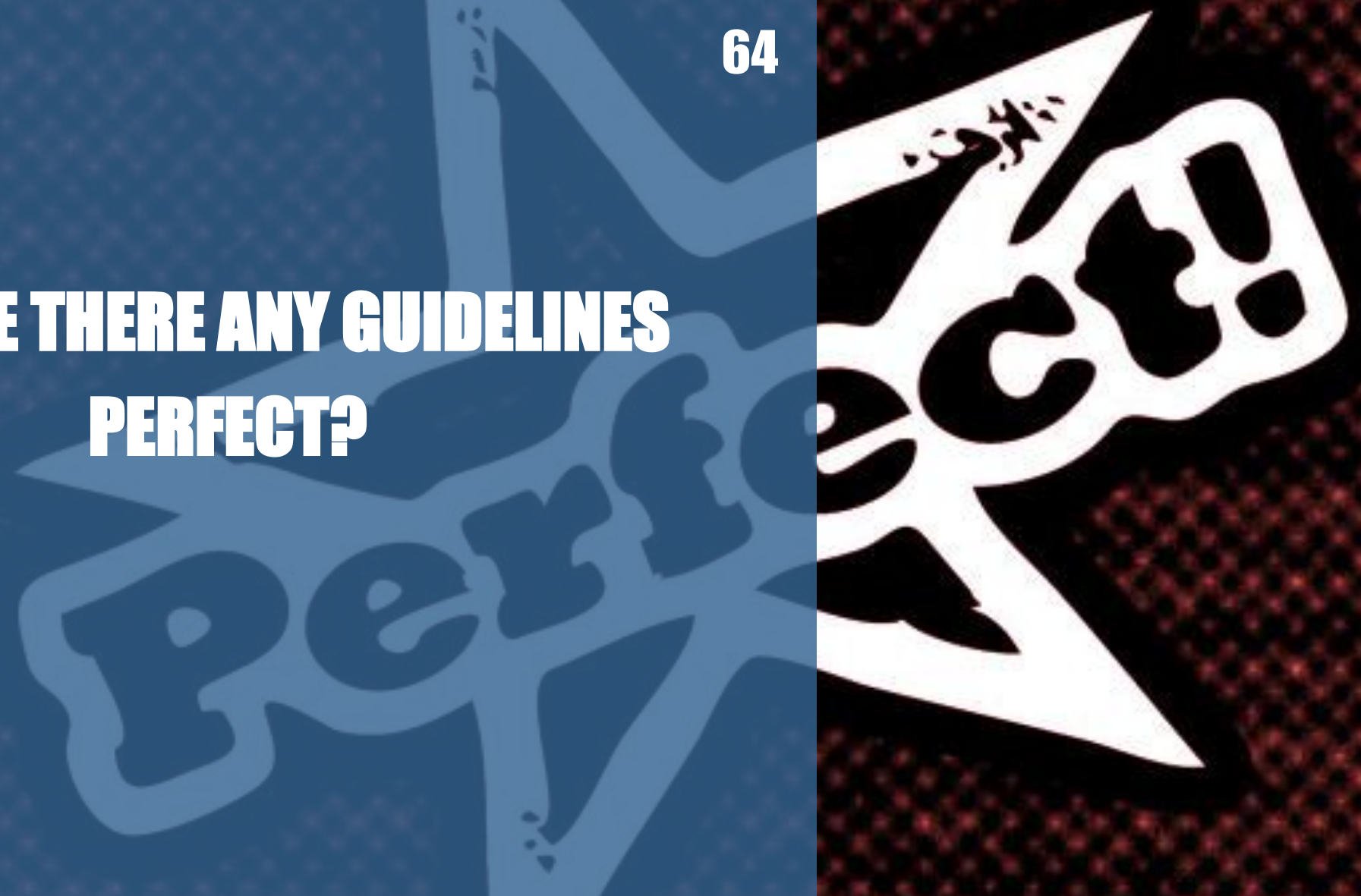
Future recommendations

1. More face to face meetings and training in guidelines development to speed up the process
2. Having unpaid people working out of hours from different nations brought its own challenges - funding for future guidelines
3. Training to members of the SOC guidelines development
4. Systematic recording of grading



**SOC8
guidelines are
a significant
improvement
from previous
SOC but are
not perfect....**

**BUT ARE THERE ANY GUIDELINES
PERFECT?**



THANKS TO

- SOC8 COMMITTEE CHAIRS AND MEMBERS
- JOHN HOPKINS UNIVERSITY AND THEIR TEAM, USA
- JUN XIA- DIRECTOR OF EVIDENCE SYNTHESIS UNIT IN SRS - UNIVERSITY OF NOTTINGHAM & KINGS COLLEGE LONDON, UK

