Analysing and Evaluating Evidence in the Context of Renal Palliative Care

Dr Fliss Murtagh FRCP PhD MRCGP MD
Cicely Saunders Institute
King's College London

Disclosure of Interests

Clinical Academic at the Cicely Saunders Institute, King's College London. Research and service development grants:

- National Institute for Health Research, UK – research grants:
  - C-CHANGE: RP-PG-1210-12015
  - C-CHANGE development: RP-DG-0709-10162
  - Non cancer place of death: SDO-FR-08-1813-257
- NHS Kidney Care, Department of Health, UK – previous service development grant:
  - The Advanced Renal Care project (2010-12)

Objectives of this talk

- What are the **general challenges** in synthesizing evidence/developing consensus statements and/or guidelines, and what are the **specific challenges for us in CKD**?
- What **level of evidence is acceptable**? How much **CKD-specific evidence** do we need and can we extrapolate from other advanced diseases?
- What **approaches to the challenges** can we utilize during this conference?
- How can we balance **working with the limitations of the evidence** against **continuing the status quo** (i.e. the huge extent of the problem and the negative consequences if care and outcomes are not improved)

General challenges

- (*Understanding the nature of the phenomenon to which the evidence relates*)
- Finding and synthesizing evidence
- Assessing the quality of the evidence
- Grading the strength of the evidence
- Deciding what is missing
- Making recommendations - moving forward

Specific challenges in advanced CKD

- (*Understanding the nature of the phenomenon to which the evidence relates*)
  1. **An ill population**
  2. **Often complex interventions**
- Assessing the quality of the evidence and grading it
  3. **Do we need different frameworks for assessment of evidence and different criteria to apply?**

1. An ill population

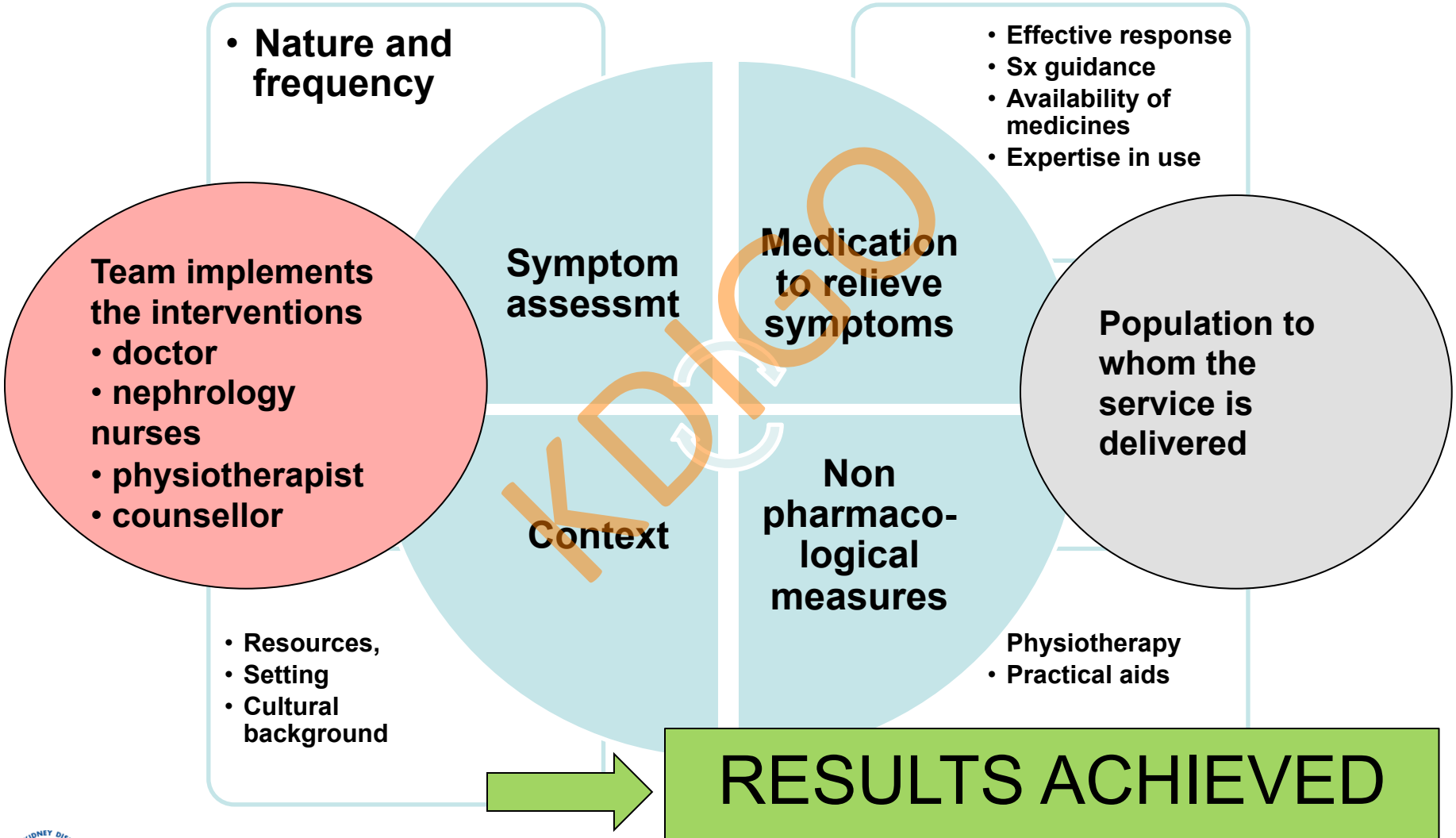
- Hard to recruit into studies
- Attrition through illness and through death
- Evidence tends to be generalised from the those at the fitter and less frail end of the CKD spectrum
- Including patients and family perspectives in building the evidence base needs time, sensitivity, and considerable ‘critical mass’

2. Usually complex interventions

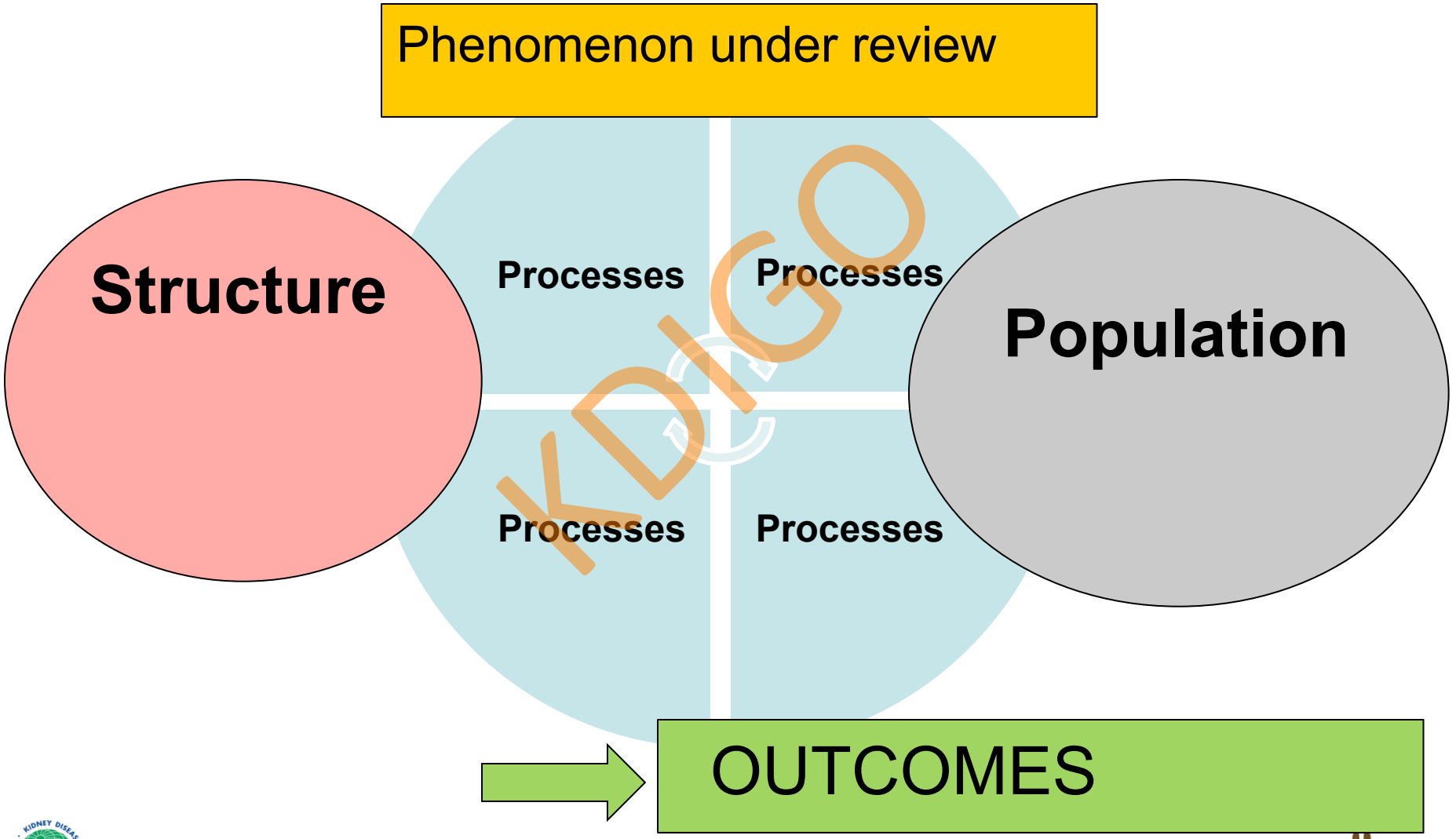
Consider nature of interventions are needed in renal palliative care. May be:

- relatively **simple** e.g. prescription to relieve uremic itch
- very **complex** e.g. implementing an integrated service to identify, assess and meet the symptom, psychological, practical and advance care planning needs of people with advanced CKD

Model of an intervention to address symptoms



Simplified model of an intervention



What makes an intervention complex?

- Number of/interactions between components within the experimental (and control) interventions
- Number and difficulty of behaviours required by those delivering or receiving the intervention
- Number of groups or organisational levels targeted by the intervention
- Number and variability of outcomes
- Degree of tailoring of the intervention required to the local setting (including cultural context)

Why does the complexity of intervention matter?

- Harder to understand and evaluate
- Much harder to build robust evidence
  - what components work and why?
- Needs different methods:
  - careful modelling and description of intervention
  - different research designs
  - usually multiple research outcomes
- Issues of generalizability
  - Local adaptation is critical, but only possible if nature/interaction of components understood

3. Assessing the evidence – should our approach be different in CKD?

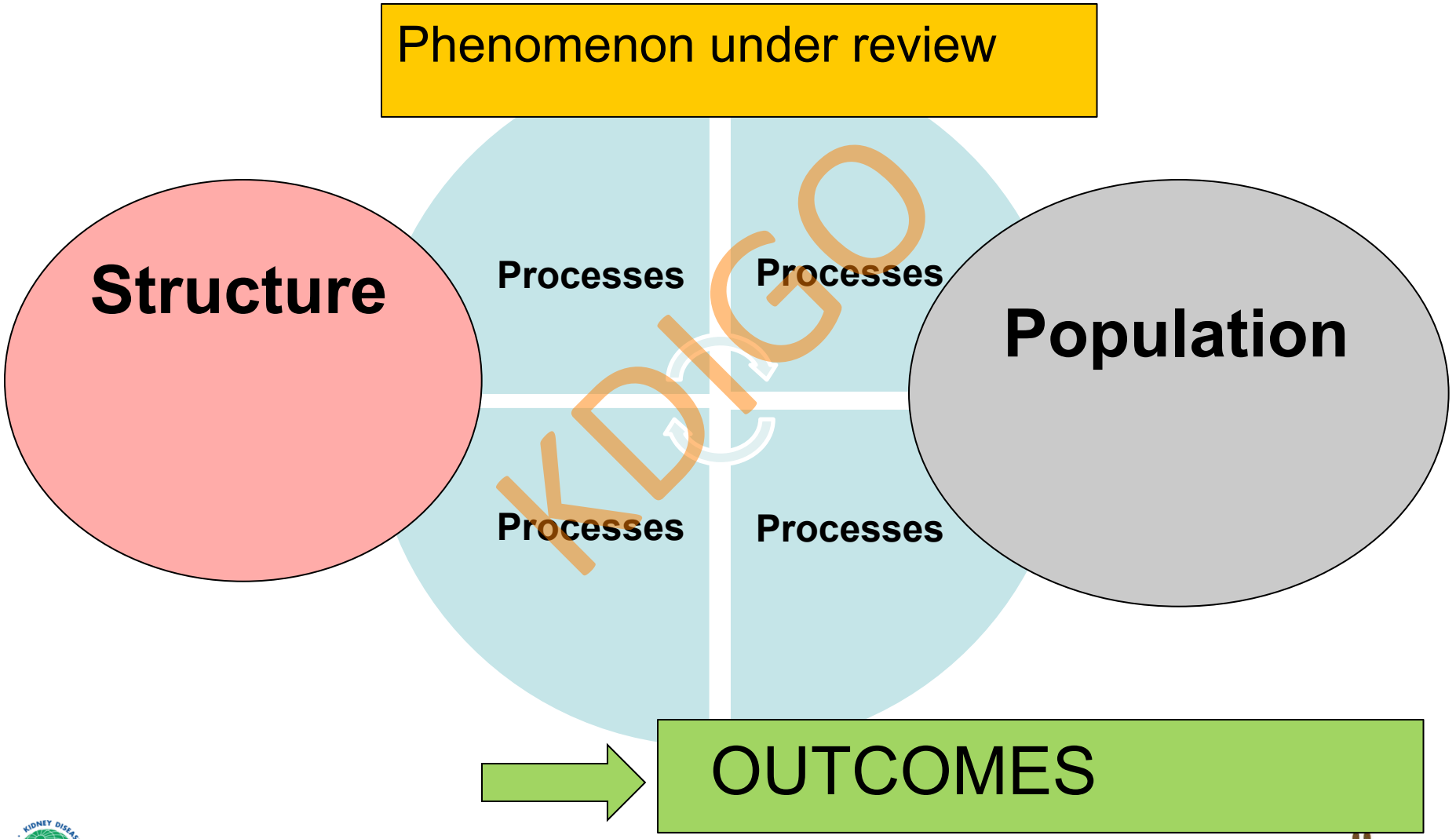
- **Yes**, because of the frailer and sicker population – harder to build evidence & greater need for improved care & outcomes
- **Yes**, because of the questions we are asking:
  - Not ***just*** about effectiveness of pharmacological interventions
  - Also about how symptoms & quality of life can be improved, prognosis communicated, dialysis decisions made, patient/family views understood and improved
- **Yes**, because of the inevitable heterogeneity of the available evidence (diverse lenses)

So how can we approach the evidence?

1. What type of question are we considering?

- Is it about effectiveness of a single component? (Is this a discreet and well defined? Are conclusions likely to be easily generalizable?)
- Or is it about effectiveness of a complex intervention ? Hard to define: e.g. a whole model of care, a decision-making intervention, advance care planning, etc
- But is it purely descriptive? Telling us about the phenomenon?
- Or may it is it epidemiological? Describing the population or factors within the population.

Simplified model of an intervention



So how can we approach the evidence?

2. Clinical applicability and importance of Q?

3. What is the evidence telling us?

- If about effectiveness
  - Is the intervention characterised well?
  - What are the outcomes – are they the right ones?
  - How closely can they be related to the intervention?
- If purely descriptive? What phenomenon described?
- If epidemiological? Which population and what factors?

So how can we approach the evidence?

4. **How can we systematically and transparently capture the quality and strength of the evidence?**
5. **Where are the gaps?**
6. **Is the evidence strong enough to move forward**

- **to a statement?**
- **to a recommendation?**
- **to consider guideline development?**

Traditional evidence classifications

- Rely very heavily on research design regardless of quality
  - Small underpowered RCT carries more weight than large 10-year longitudinal study
- RCTs - 'Gold standard' for research design
  - Not suitable to answer all questions – mainly effectiveness
  - Do provide overall evidence of effectiveness (design addresses bias)
  - May or may not be clinical relevant

Evidence typically considered

- Based on
  - **P**atients – which people/populations?
  - **I**ntervention – what interventions (and their characteristics)?
  - **C**omparator/Control
  - **O**utcomes – which are relevant?

► Search strategy based on same criteria

What approach should we adopt?

(Baker 2010)

- SIGN – Scottish Intercollegiate Guidelines Network
- GRADE - Grading of Recommendations Assessment, Development and Evaluation
- GATE – the Graphic Appraisal Tool for Epidemiology
- **NSF- LTC – the UK National Service Framework for Long Term Conditions**

UK NSF for Long Term Conditions

- Recognises two main groups of evidence
  - Expert evidence
    - Opinion based on experience
      - E1: Users/carers
      - E2: Professionals or other stakeholders
  - Research evidence
    - Categorised by 3 criteria
      - Design
      - Quality
      - Applicability
        - » Direct – evidence from relevant population
        - » Indirect – evidence from other conditions

Classification of research design

	Primary Research-based Evidence
P1	**Quantitative research** (e.g. experimental, longitudinal, cross-sectional)
P2	**Qualitative research delivering descriptive data** (e.g. focus group, rich narrative consensus methodologies)
P3	**Mixed methods (qualitative and quantitative)** (e.g. a survey using both standardised measurement tools and detailed interviews)
	Secondary Research-based Evidence
S1	**Meta-analysis of existing studies**
S2	**Secondary analysis of existing data**
	Review-based Evidence
R1	**Systematic review of existing research**
R2	**Descriptive or summary reviews of existing research**

Quality rating

Quality assessment item	Score
Are research question/aims and design clearly stated?	
Is the research design appropriate? for stated aims and objectives of research	
Are the methods clearly described?	
Is the data adequate? to support the authors' conclusions / interpretations	
Are the results generalisable?	
out of possible Total Score	**10**

Each quality item rated: "Yes" = 2, "In part" = 1, "No" = 0

Grade of recommendation

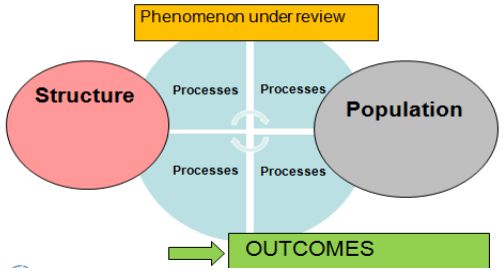
- Research Grade A
  - >1 study of high quality score (≥7/10) *and*
  - At least one of these "Direct" applicability
- Research Grade B
  - 1 high quality study *or*
  - >1 study of medium quality (4-6/10) *and*
  - At least one of these "Direct" applicability
- Research Grade C
  - 1 medium quality study *or*
  - Lower quality studies only (grade 2-3/10)
  - Indirect studies only

For quality & strength of evidence, this gives:

- Each quoted reference
  - Carries a category (E/P/S/R) and a number (1,2 or 3). Note can be two e.g. E1+2
- Each evidence-based statement
  - Carries a grade based on quality and strength
    - Expert, Research A, Research B, or Research C
- *P1 High Direct* (meaning a high quality quantitative study of direct applicability)
- *E1+2* (meaning expert opinion from both patients/families and professionals)

Questions to consider

1. What type of question are we considering?
2. How clinically relevant and important?
3. What is the evidence telling us?
4. Where are the gaps?
5. What quality and strength of evidence? (perhaps using NSF-LTC criteria)
6. Where next – statement, recommendation, or guideline development?



Thank you

- Cicely Saunders Institute website
  www.csi.kcl.ac.uk/

- fliss.murtagh@kcl.ac.uk

Useful resources

The following section contains definitions and signposts relevant web resources

Clinical guidelines - definition

"*Systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances*"

Attributes of good guidelines: validity, reliability, clinical applicability, clinical flexibility, clarity, multidisciplinary process, scheduled review, and documentation

Institute of Medicine Committee to Advise the
Public Health Service on Clinical Practice Guidelines.
Clinical practice guidelines: directions for a new program. 1990

Guidelines International Network

- Global network of 86 organisations and 107 individual members representing 45 countries
- Supports evidence based health care and improved health outcome by reducing inappropriate variation throughout the world
- Annual conferences

www.g-i-n.net

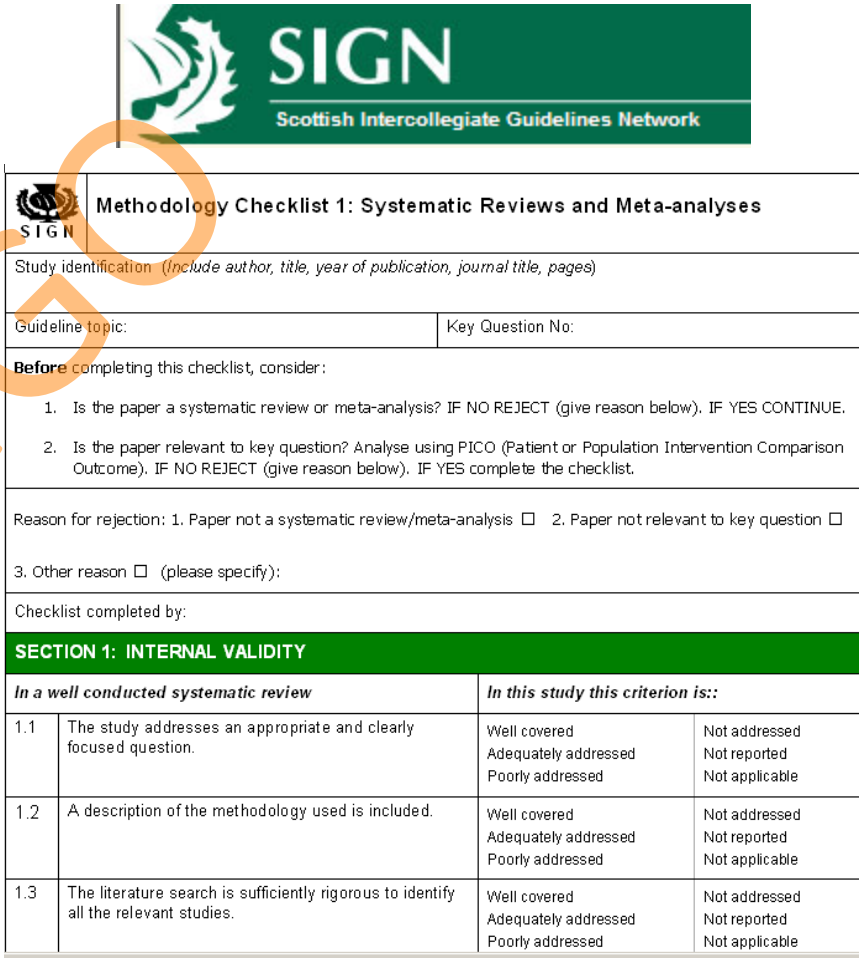
Scottish Intercollegiate Guidelines Network (SIGN)

- Develops evidence based practice guidelines for the NHS in Scotland
- Useful resources for guideline methdology
- “Guideline development in 50 easy steps“
- Methodology checklists for critical appraisal of different study designs
- Online tutorials
- Guideline development handbook

www.sign.ac.uk

SIGN: Five methodology checklists

1. Systematic reviews and meta-analyses
2. Randomised controlled trials
3. Cohort studies
4. Case control studies
5. Diagnostic studies



SIGN
Scottish Intercollegiate Guidelines Network

Methodology Checklist 1: Systematic Reviews and Meta-analyses

Study identification (*Include author, title, year of publication, journal title, pages*)

Guideline topic:	Key Question No:

Before completing this checklist, consider:

1. Is the paper a systematic review or meta-analysis? IF NO REJECT (give reason below). IF YES CONTINUE.
2. Is the paper relevant to key question? Analyse using PICO (Patient or Population Intervention Comparison Outcome). IF NO REJECT (give reason below). IF YES complete the checklist.

Reason for rejection: 1. Paper not a systematic review/meta-analysis ☐ 2. Paper not relevant to key question ☐

3. Other reason ☐ (please specify):

Checklist completed by:

SECTION 1: INTERNAL VALIDITY

In a well conducted systematic review		*In this study this criterion is::*	
1.1	The study addresses an appropriate and clearly focused question.	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
1.2	A description of the methodology used is included.	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
1.3	The literature search is sufficiently rigorous to identify all the relevant studies.	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable

SIGN: Levels of evidence

LEVELS OF EVIDENCE

1++	High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias
1+	Well conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias
1-	Meta-analyses, systematic reviews, or RCTs with a high risk of bias
2++	High quality systematic reviews of case control or cohort studies High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
2+	Well conducted case control or cohort studies with a low risk of confounding or bias and a moderate probability that the relationship is causal
2-	Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal
3	Non-analytic studies, eg case reports, case series
4	Expert opinion

GRADES OF RECOMMENDATION

Note: The grade of recommendation relates to the strength of the evidence on which the recommendation is based. It does not reflect the clinical importance of the recommendation.

A	At least one meta-analysis, systematic review, or RCT rated as 1++, and directly applicable to the target population; or A body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results
B	A body of evidence including studies rated as 2++, directly applicable to the target population, and demonstrating overall consistency of results; *or* Extrapolated evidence from studies rated as 1++ or 1+
C	A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; *or* Extrapolated evidence from studies rated as 2++
D	Evidence level 3 or 4; or Extrapolated evidence from studies rated as 2+

GRADE

- Grading of Recommendations Assessment, Development and Evaluation (short GRADE) Working Group
- Widely accepted framework for assessing the evidence and making recommendations
- www.gradeworkinggroup.org/index.htm

GRADE

- The Grading of Recommendations Assessment, Development and Evaluation (short GRADE) Working Group
- created in response to the need for a more unified and transparent approach to guidelines creation and reporting
- adopted as the standard for guideline development by over 50 international organizations (WHO), the Cochrane Collaboration, SIGN, AHRQ, and the Centers for Disease Control and Prevention)
- Need training in using it

Canfield et al World J Urol 2011

GRADE

- GRADE system suggests attempting to identify all potentially relevant outcomes for each specific question and rate their relative importance a priori

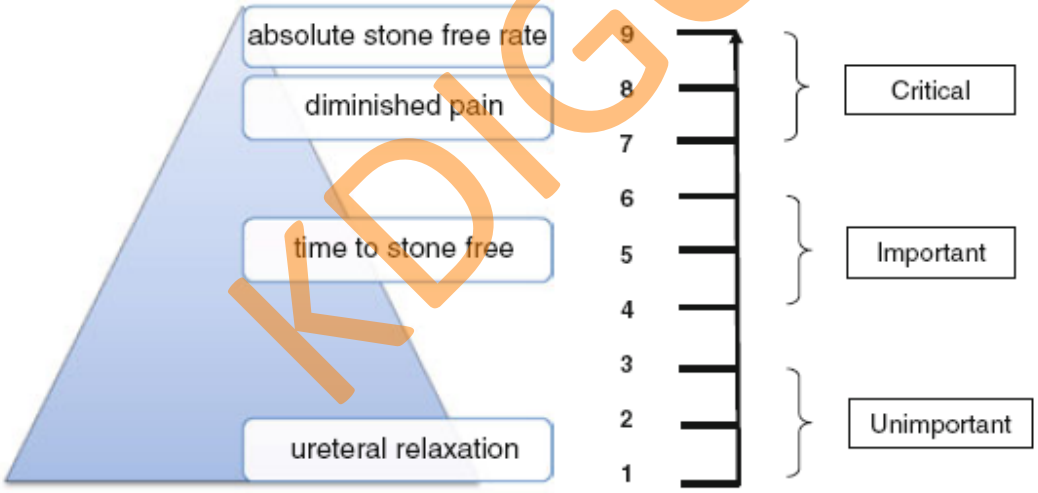


Fig. 1 Example of potential hierarchy of outcomes for medical expulsive therapy

Canfield et al World J Urol 2011

GRADE evidence rating

- Outcome specific
- As quality of evidence might vary across outcomes

Table 3 Definitions for GRADE quality ratings and strength of recommendations

Quality of evidence		Strength of recommendation
High quality	++++	Strong recommendation for using an intervention
Moderate quality	+++	Weak recommendation for using an intervention
Low quality	++	Weak recommendation against using an intervention
Very low quality	+	Strong recommendation against using an intervention

Canfield et al World J Urol 2011

GATE

- Graphic Appraisal Tool for Epidemiology (GATE) is a visual framework that illustrates the generic design of all epidemiologic studies
- It was designed by the School of Population Health, University of Auckland, New Zealand
- Available at:

  http://www.fmhs.auckland.ac.nz/soph/depts/epi/epiq/_docs/gateframe.pdf

Standards for guidelines

- Appraisal of Guidelines Research & Evaluation (AGREE)
  - Guideline appraisal tool (www.agreetrust.org)
  - Checklist to assess quality of guideline
  - Used by international guideline bodies