

# How much do experts' opinions matter when data is sparse? Value of expert elicitation methods in addressing COVID-19 pandemic challenges

---

Virtual ISPOR 2021

May 17-20, 2021

PRECISION**heor**



# Disclosures and Acknowledgements

---

- Phani Veeranki is an employee of PRECISIONheor, which provides health economics and outcomes research consulting services to life sciences companies.
- Tony O'Hagan is a statistical consultant who advises organizations, including several pharmaceutical companies, on elicitation of expert knowledge.
- Laura Bojke is a member of the technology appraisal committee for the National Institute for Health and Care Excellence (NICE)

# Panelists' Introduction

---

PRECISION**heor**



# Panelists

---



## **Phani Veeranki, MD, DrPH**

Dr. Veeranki is a Director at PRECISIONheor, Los Angeles, CA. He is a principal investigator of several industry and non-industry funded projects that generate real-world evidence to support novel healthcare innovations. He is an expert in expert elicitation methods and applies the methodology in multiple research projects.



## **Anthony O'Hagan, PhD**

Professor O'Hagan is Emeritus Professor of Statistics at the University of Sheffield, UK. He is a co-developer of the Sheffield Elicitation Framework (SHELF) and an expert in the methodology and applications of Bayesian statistics.



## **Laura Bojke, PhD**

Laura Bojke is a Professor of Health Economics at the Centre for Health Economics, University of York. She is a member of an assessment group for NICE and an appraisal committee member. She is the co-lead of a project to determine the appropriateness of expert elicitation methods for health technology assessment.



## **Donna Fick, RN, PhD, FAAN**

Dr. Ficks is the Elouise Ross Eberly Professor of Nursing at PennState College of Nursing. She is an expert in Delphi methodology and has applied the methodology in several geriatric projects including development of 2019 AGS Beers Criteria

# Despite many emerging data on COVID-19, there still remains a gap

## Emerging data since past 14 months

### COVID-19 Databases and Journals

Updated September 2, 2020

[COVID-19 Research Guide Home](#)

Below are selected databases and journals to help researchers find scholarly articles about COVID-19 (2019 Novel Coronavirus).

### Risk Factors Associated With In-Hospital Mortality in a US National Sample of Patients With COVID-19

Ning Rosenthal<sup>1</sup>, Zhun Cao<sup>1</sup>, Jake Gundrum<sup>1</sup>, Jim Sianis<sup>1</sup>, Stella Safo<sup>1</sup>

### Epidemiology of COVID-19

[Sudipta Dhar Chowdhury](#)<sup>1</sup> and [Anu Mary Oommen](#)<sup>2</sup>

### Healthcare disparities among anticoagulation therapies for severe COVID-19 patients in the multi-site VIRUS registry.

Kirkup C, Pawlowski C, Puranik A, Conrad I, O'Horo JC, Gomaa D, Banner-Goodspeed VM, Mosier JM, Zabolotskikh IB, Daugherty SK, Bernstein MA, Zaren HA, Bansal V, Pickering B, Badley AD, Kashyap R, Venkatakrishnan AJ, Soundararajan V.

## Situations where data is inadequate

### Guidance on nebulization during the current COVID-19 pandemic

[Mario Cazzola](#),<sup>a,\*</sup> [Josuel Ora](#),<sup>b</sup> [Andrea Bianco](#),<sup>c</sup> [Paola Rogliani](#),<sup>a,b</sup> and [Maria Gabriella Matera](#)<sup>d</sup>

### The Impact of COVID-19 on Outpatient Visits in 2020: Visits

### Clinical Issues & Guidance for Elective Surgery

--Several professional organizations and societies have re-evaluated their strategies in clinical practice to minimize virus transmission, despite **lack of data**.

-Such situations of data inadequacies are also noticed in rare and orphan diseases.

Potential solution for situations with inadequate data are **expert elicitation methods**

# Agenda

---

- |   |  |                                 |
|---|--|---------------------------------|
| 1 | Expert Elicitation Methods (SHELF) in Quantifying Expert Opinions                          | Prof. Anthony O'Hagan           |
| 2 | Expert Elicitation Methods (Delphi) to Inform Research and Clinical Decision-Making        | Prof. Donna Fick                |
| 3 | Expert-elicitation methods to Inform Health Technology Assessments and Formulary Decisions | Prof. Laura Bojke               |
| 4 | Moderated Discussion   | Moderated by Dr. Phani Veeranki |
-

# Expert Elicitation Methods (SHELF) in Quantifying Expert Opinions

---

PRECISION**heor**



# Expert Knowledge as Decision Support

When decisions must be made in rapidly evolving contexts, data are typically sparse

## Expert knowledge becomes **important and valuable**

- e.g. How transmissible will this new variant of the SARS-COV-2 virus be?
- **Experts are used to providing input to such questions, but often only qualitatively**
  - e.g. “I think it will be about as transmissible as (or much more transmissible than) the Brazilian variant”
- **Quantitative input is much more valuable**
  - e.g. “I estimate the  $R_0$  for this variant to be 4.5”
- **Acknowledging that there is (typically substantial) uncertainty**
  - e.g. “There is an 80% chance that  $R_0$  will be between 2.5 and 6”
- **Judgements are typically sought from multiple experts**
  - What the decision-maker really needs is an expression of their **combined** knowledge

# Elicitation

---

- **Expert Knowledge Elicitation is the process of:**

- Representing the knowledge of one or more persons (experts) concerning an uncertain quantity
- Presented as a **probability distribution** for that quantity

- **Typically conducted as a dialogue between:**

- The experts – who have substantive knowledge about the quantity (or quantities) of interest – and
- A facilitator – who has expertise in the process of elicitation

Ideally, it will be conducted face-to-face, but may also be done by video-conference

# The importance of the facilitator

---

- **The elicited probability distributions must meet the standards of good science**
  - Must be carefully considered
  - Take in full account of the available evidence
  - Should not overstate nor understate the uncertainty
- **But research has revealed many ways in which experts make poor judgements**
  - Particularly, when questioned in naïve ways
- **Judgements are influenced, and potentially biased, by**
  - How questions are phrased
  - The order in which they are asked
  - Interactions between experts
- **The role of the facilitator is crucial**
  - Must follow a well-constructed **protocol** to minimise the potential for bias
  - Ensure all opinions are given due weight and consideration

# The SHELF protocol

---

## ■ Preparation

- Assemble a **dossier** of available evidence
- Recruit experts and get their commitment to the project
- Train the experts in the kinds of **probabilistic judgements** they will be required to make

## ■ Facilitated elicitation workshop

- Usually **4 to 8** experts
- Conduct practice runs to familiarise experts with the process and to reinforce preparatory training
- Make private individual judgements first to establish each expert's initial position
- Be involved in group discussions to understand differences in their individual judgements
- View group judgements from the perspective of a Rational Impartial Observer (RIO)
- Feedback and validity checking throughout

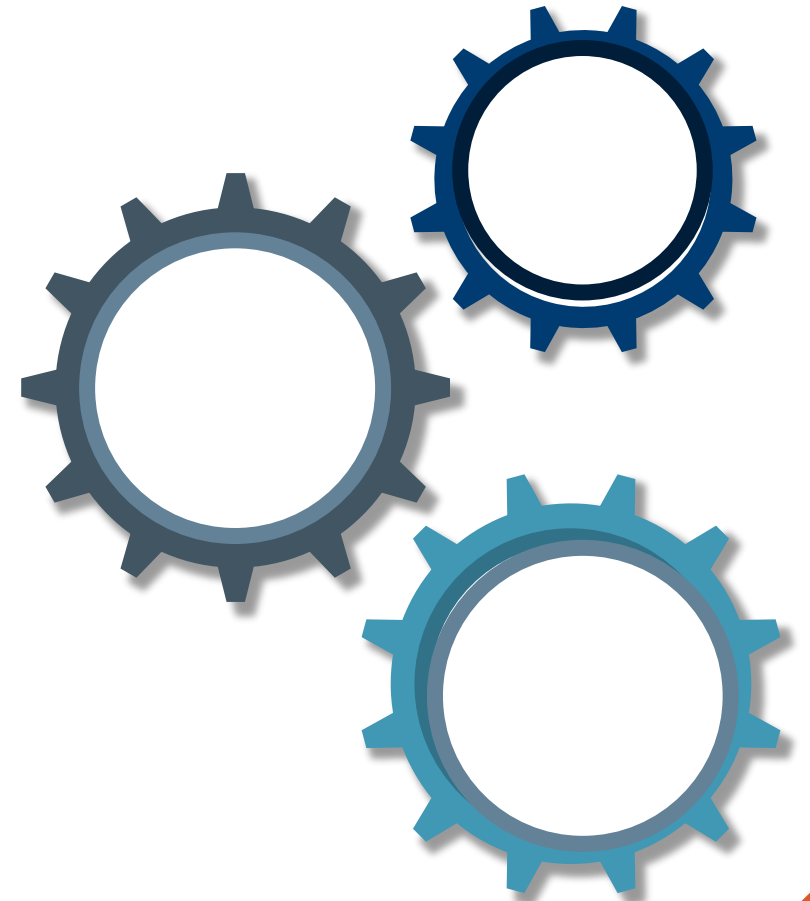
## ■ Reporting

- SHELF templates provide **auditable record** of the process

# Summary

---

- **Expert judgement is an important input to decision-making**
  - Particularly in rapidly-changing, data-poor contexts like emerging threats, pandemics, etc.
- **Expert judgements need to be elicited with great care**
  - Poor judgements are worse than useless
  - Elicitation is not an easy option
- **Elicitation should be conducted by a skilled facilitator**
  - Following a recognised protocol such as SHELF
  - Should be fully-documented



# Expert Elicitation Methods (Delphi) to Inform Research and Clinical Decision-Making

---

Donna Fick, RN, PhD, FAAN

PRECISION**heor**



# Use of Delphi and Modified Delphi Technique For Expert Elicitation

---

- Delphi is a research design, used to find solutions or answers to problems that are harder to solve or where data is missing. It combines a qualitative and quantitative approach as a solution.
- History of Delphi goes back thousands of years but most in the field think of the “Classic Delphi” from the 1960s (Helmer, 1967). This is where Delphi was used for consensus in complex military issues. It is most often thought of for future forecasting but has been used more broadly over time.

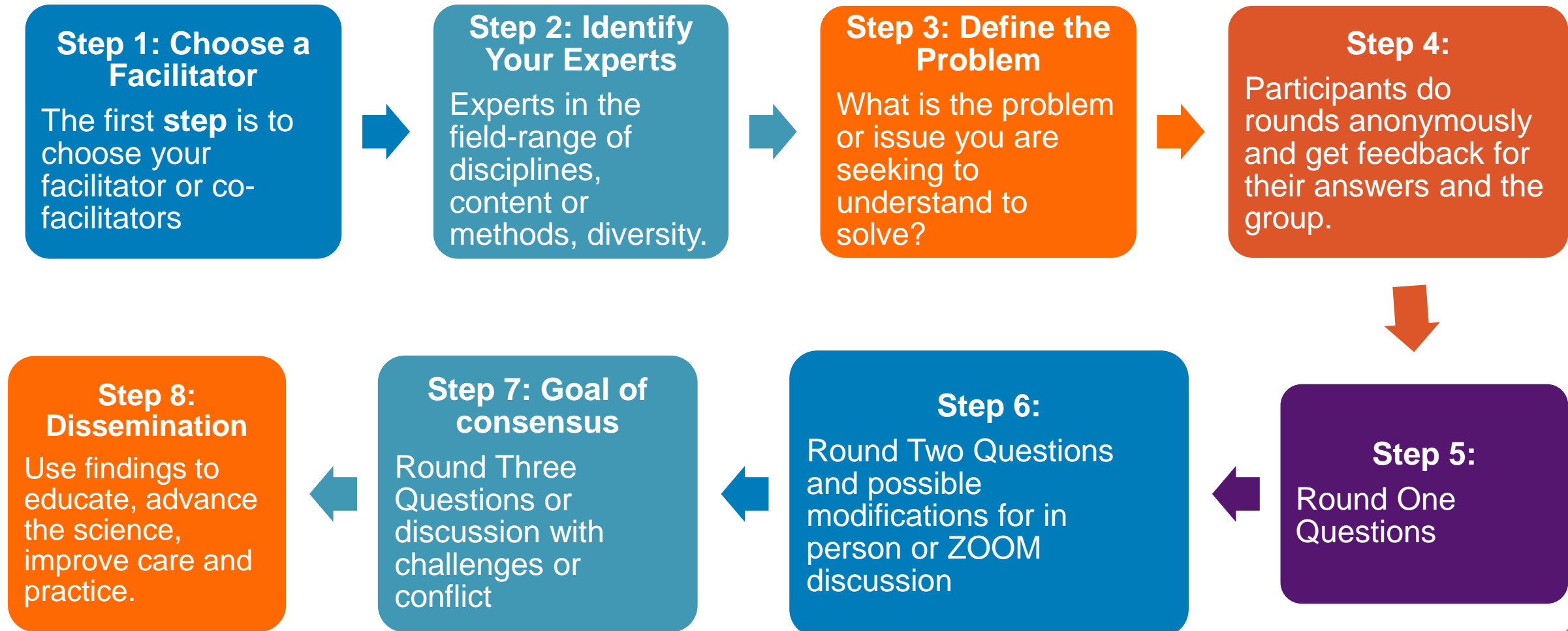
# When do we use Delphi in clinical research and clinical decisions?



- In problems or populations where little data exists
- The problem is poorly understood and subjective judgements by clinicians from their experience may help inform the problem
- Times when anonymous communication is helpful, or money or time make this more practical
- Rare diseases, a pandemic, or under-represented populations
- Where expert opinion can add to understanding of very challenging or new clinical issues in the field
- Where there is data, but still clinical decision-making needed for decisions or development into practice

# Use of Delphi and Modified Delphi Technique

## Delphi Technique: A Step-by-Step Guide



# 2019 AGS BEERS Criteria

From THE AMERICAN GERIATRICS SOCIETY

## A POCKET GUIDE TO THE 2019 AGS BEERS CRITERIA®

This guide has been developed as a tool to assist healthcare providers in improving medication safety in older adults. The role of this guide is to *inform* clinical decision-making, research, training, quality measures and regulations concerning the prescribing of medications for older adults to improve safety and quality of care. It is based on *The 2019 AGS Beers Criteria® for Potentially Inappropriate Medication Use in Older Adults*.

Originally conceived of in 1991 by the late Mark Beers, MD, a geriatrician, the Beers Criteria catalogues medications that cause side effects in older adults due to the physiologic changes of aging. In 2011, the AGS sponsored its first update of the criteria, assembling a team of experts and using an enhanced, evidence-based methodology. Since 2011, the AGS has been the steward of the criteria and has produced updates using an evidence-based methodology and rating each Criterion (quality of evidence and strength of evidence) using the American College of Physicians' Guideline Grading System, which is based on the GRADE scheme developed by Guyatt et al.

The full document, along with accompanying resources, can be found in its entirety online at [geriatriccareonline.org](http://geriatriccareonline.org).

### INTENDED USE

The goal of this guide is to improve care of older adults by reducing their exposure to Potentially Inappropriate Medications (PIMs).

- This should be viewed as a guideline for identifying medications for which the risks of their use in older adults outweigh the benefits.
- These criteria are not meant to be applied in a punitive manner.
- This list is not meant to supersede clinical judgment or an individual patient's values and needs. Prescribing and managing disease conditions should be individualized and involve shared decision-making.
- These criteria also underscore the importance of using a team approach to prescribing and the use of non-pharmacological approaches and of having economic and organizational incentives for this type of model.
- A companion piece that addresses the best way for patients, providers, and health systems to use (and not use) the AGS Beers Criteria® was also developed. The document can be found on [geriatriccareonline.org](http://geriatriccareonline.org).

The criteria are not applicable in all circumstances (i.e. patients receiving palliative and hospice care). If a provider is not able to find an alternative and chooses to continue to use a drug on this list in an individual patient, designation of the medication as potentially inappropriate can serve as a reminder for close monitoring so that adverse drug effects can be incorporated into the electronic health record and prevented or detected early.

## Example using Delphi where issues poorly understood: Delirium Superimposed on Dementia (COVID symptom)

---

**The Question:** How to define and measure delirium (acute, reversible confusion) and delirium severity in a person with dementia?



**Panelists/Experts:** Geriatricians, Direct Care Nurse, DSD Clinical Scientist, Neuropsychologist, Neurologist, Psychiatrist

**Challenges:** Overlapping symptoms, poor understanding of phenotypes and acute changes

## Example During Height of COVID-19

**Challenge:** How to minimize medication use and caregiver and resident contact during an infectious disease pandemic?

# Stepwise approach

BOX:

## Stepwise approach to prioritizing and implementing recommendations

Consider a stepwise approach for implementing recommendations in your community. A suggested approach is listed below, although adaptation is encouraged based on local circumstances.

### Priority for implementation

#### 1. Changes that are essential for infection control

- Transition from nebulizers to hand-held inhalers for residents needing inhaled therapy when feasible, safe, and available. (Table 4)
- Among residents with known or suspected COVID-19: changes that reduce frequency, duration, and infection risk of medication passes (All tables)

#### 2. Changes that are generally low risk, can be quickly evaluated for individual appropriateness, and can be done immediately.

- Discontinuation of medications that do not provide benefit for most residents, can be stopped abruptly, and do not need extensive monitoring after discontinuation (Table 1)
- Change from short- to longer-acting medications where conversion is routine and changes are typically well-tolerated (Table 1)
- Conversion to dosing forms that are easier to use and administer (Table 2)
- Consolidate and liberalize administration times for medications that do not need to be given at very specific times or intervals (Table 3)
- Enhanced hygiene measures during medication passes (Table 4)

#### 3. Changes that are generally low risk but may take more time for implementation, individual evaluation, communication with care team and resident, and monitoring

- Changes in insulins, analgesic regimens (Tables 1 and 2)

# Implications and summary

---

- An obvious limitation of the traditional Delphi method is that it does not elicit the experts' uncertainty. Therefore, a version known as **Probabilistic Delphi** can be used to make the same set of judgements as in a SHELF method. They should make these judgements in the same sequence, beginning with credible limits. The facilitator gathers all the respondents' judgements, and in sending out the next questionnaire, provides feedback summaries of those judgements. This can be used for those that are less important and do not need the full SHELF.
- The methods should be transparent, rigorous, and replicable.
- Delphi has limitations and challenges when used in a clinical research context such as what level of panel interaction to have, what level of evidence review, and context setting to the panel about the problem, how to choose panelists (diversity, range of experts on topic), etc.

# Implications and summary (continued)

- Expert elicitation can be very useful as a tool and method for clinical research and making clinical decisions
- Delphi can be adapted, but an important part of Delphi is that the questionnaires be anonymous
- It can be done rapidly at a distance and is useful during times where answers are needed quickly for **NEW** problems happening in real-time, such as the COVID-19 pandemic

# References

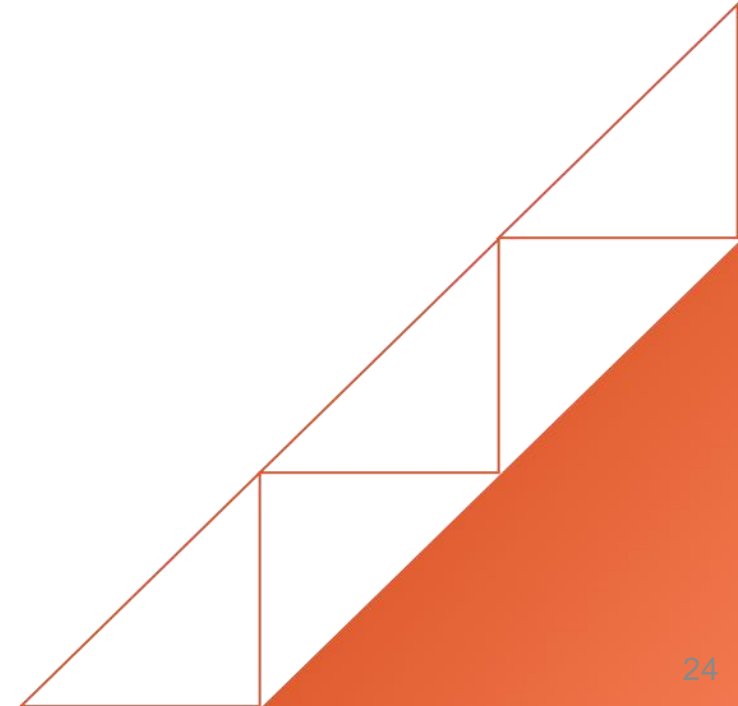
---

- Scott IA, Rigby D, Hilmer SN. Optimising medication management during the COVID- 19 pandemic. *J Pharm Pract Res.* 2020 Jun;50(3):186-9.
  - Brandt N, Chou J. Optimizing Medication Management During the COVID-19 Pandemic: It Takes a Village. *Journal of Gerontological Nursing.* 2020;46(7):3-8.
  - Brandt N, Steinman MA. Optimizing Medication Management During the COVID ,Äê19 Pandemic: An Implementation Guide for Post-Acute and Long-Term Care. *J Am Geriatr Soc.* 2020 Jul;68(7):1362-5.
  - American Geriatrics Society Beers Criteria® Update Expert Panel, Fick DM, Semla TP, Steinman M, Beizer J, Brandt N, Dombrowski R, DuBeau CE, Pezzullo L, Epplin JJ, Flanagan N, Morden E, Hanlon J, Hollmann P, Laird R, Linnebur S, Sandhu S. (2019). American Geriatrics Society 2019 Updated AGS Beers Criteria® for potentially inappropriate medication use in older adults. *J Am Geriatr Soc.* 67(4):674-694. <https://doi.org/10.1111/jgs.15767>.
-

# Expert-elicitation methods to inform health technology assessments and formulary decisions

---

PRECISION**heor**



# Evidence requirements for HTA and formulary decisions

---

- **Such decisions require multiple types of evidence**
  - Treatment effects, costs, quality of life, natural history progression, disease incidence
- **Randomized controlled experiments (generally) prefer to estimate a treatment effect**
  - Non-RCT evidence often used for other elements required to make decisions, e.g., costs
- **RCT evidence is often uncertain**
  - Evidence may not be on 'final' outcomes, e.g., cancer products licensed on evidence of progression-free survival
  - Evidence base may not be well developed, e.g., early access to medicines scheme
- **Expert opinions are often used in HTA at various stages of the process**
  - Opinions given on the interpretation of the evidence
  - Formal elicited judgements part of evidence pack

# Use of expert elicitation across HTA organisations

---

- **14 country-specific**, HTA methods guides were reviewed
  - Most HTA organisations mention that expert opinion may be used to inform multiple types of parameters, such as efficacy, HRQoL, costs, resource use, and adverse events
- These guidance consistently suggest that evidence elicited from experts is of low quality and recommend its use only in the absence of other (empirical) evidence
- Most HTA organisations require that when evidence derived from experts is used, the **process of experts' selection and methods for elicitation is clearly described to ensure transparency**
- Only PBAC and EUnethta provide additional details on which processes are considered appropriate and how experts should be selected or how experts' beliefs are elicited and combined

Full details of this review at: [Elicitation-NICE-final-report-York\\_01042020.pdf](#)

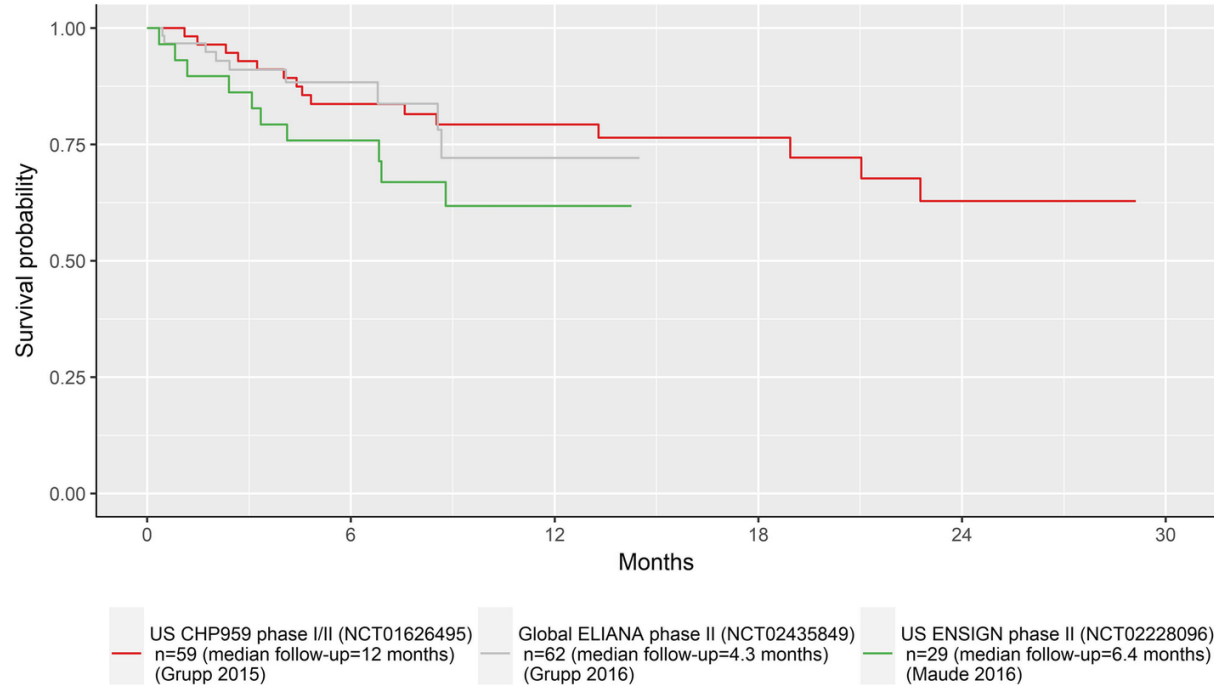
# Use of expert elicitation in NICE assessments

---

- A recent review of **25 company submissions**, appraised by NICE (STA process), found that expert judgement is ubiquitously used in company submissions **(23/25)**.
- Most commonly, the submissions referred to gathering expert judgements using informal, unstructured methods, that can be described as either **ad-hoc methods (55.1%)** or **advisory boards (38.8%)**.
- The use of expert elicitation is less common than the use of expert opinion and is generally restricted to cases where there are significant gaps in the evidence base.
- Expert opinions are often required when extrapolating evidence from one population to another or over a time-horizon unobserved in clinical studies
- NICE recommends that any extrapolation should consider ‘both clinical and biological plausibility of the inferred outcome as well as its coherence with external data sources’; however, they do not specify any methodologies to accomplish this.
  - Expert elicitation can be used in this context

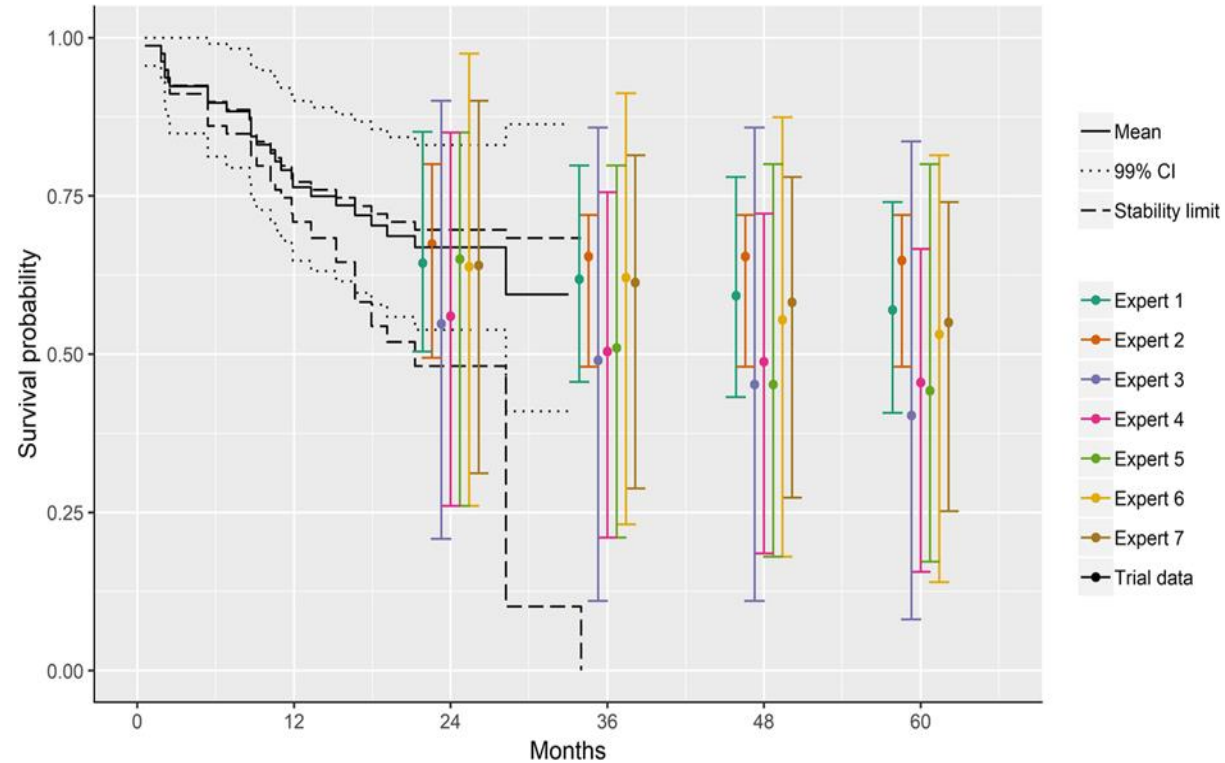
# Example from: CAR-T therapy for children and young adults with relapsed or refractory acute lymphoblastic leukemia

From: [Integrating expert opinion with clinical trial data to extrapolate long-term survival: a case study of CAR-T therapy for children and young adults with relapsed or refractory acute lymphoblastic leukemia](#) (Cope, et al, 2019)



- Tisagenlecleucel (Kymriah®) is the first CAR-T therapy approved by the FDA for the treatment of pediatric and young adult patients.
- Previous assessments of cost-effectiveness showed the reimbursement decision was highly sensitive to choice of parametric model fitted to observed data.
- Methods adapted from SHELF were used to obtain survival estimates from multiple experts, which were combined with the empirical data to estimate long-term survival curves.

# Results from CAR-T elicitation



- Extrapolated survival, incorporating expert estimates, was compared to the longer follow-up from the empirical study
- At 24 months, the expert estimates were generally **very close** to the observed survival

# Areas for consideration

---

- How should experts' estimates be used in HTA?
- How can we ensure the validity of elicited estimates?
- Are there specific areas of HTA that may be more amenable or needing of expert elicitation?
- Are there a minimum set of requirements for elicitation used to inform HTA?

# What might a minimum set of requirements look like?

---

- Efforts should be made to increase validity and transparency of expert-elicited used to inform HTA and formulary decisions
  - Decision-makers should be assured of the quality of the elicited estimates
  - As a minimum, HTA agencies should require thorough reporting of the elicitation, for example an elicitation protocol
- Expert elicitation should follow processes appropriate for that setting
  - Flexibility in approaches may be needed depending on the context, e.g. consensus tasks may be challenging in rare diseases where experts are geographically disparate, but may also be beneficial where there is little 'observed' clinical practice
- Expert elicitation in HTA and formulary decisions needs to reflect heterogeneity in who may be regarded as an expert
  - In some tasks, professions allied to medicine have more observed experience of the quantities of interest
- Where similarly quantities are likely to be elicited on a frequent basis, HTA agencies may want to consider expert panels trained in elicitation



# Moderated Discussion