



What Statistics and Analytics Resources are Available to HEOR Professionals?

Eberechukwu Onukwugha, PhD
Helene Karcher, PhD
Douglas Faries, PhD

Tuesday May 17, 2022 @ 1:30-2:30pm ET

Acknowledgement and Thanks for the ISPOR Health Economics and Outcomes Research Competencies Framework™

- Nancy Berg, Dick Willke, Betsy Lane and the ISPOR Staff
 - ISPOR Board of Directors
 - ISPOR Institutional Council
 - ISPOR Faculty Advisor Council
 - ISPOR Student Council
-
- The Competency Inventory is proprietary to ISPOR. No part of this document may be disclosed or repurposed in any manner without the prior written consent of ISPOR—The professional society for health economics and outcomes research.

Team members

- Eberechukwu Onukwugha, PhD, University of Maryland
- Doug Faries, PhD, Eli Lilly (Advisor)
- Helene Karcher, PhD, Novartis Pharma AG (Advisor)
- Tsung-Ying Lee, PhD student, University of Maryland
- Adaeze Amaefule, PharmD student , University of Maryland

Objectives of this Session

1. Describe and define the available topical content within the 'Statistics and Analytics' category.
2. Identify training tools and resources for gaining the statistical knowledge and skills necessary to conduct and report health economics and outcomes research
3. Discuss gaps in training tools and resources



Brief History of the ISPOR Competency Framework Initiative

- Initiated in 2016
- Idea originated with the Institutional Council
- Faculty Advisor Council engaged soon thereafter
- Published in 2020 in *Value in Health*
- Present findings related to the Statistical and Analytics competency
- Opportunities for developing content information for other domains



Open Access article available here:

<https://www.sciencedirect.com/science/article/pii/S1098301520321240>



1

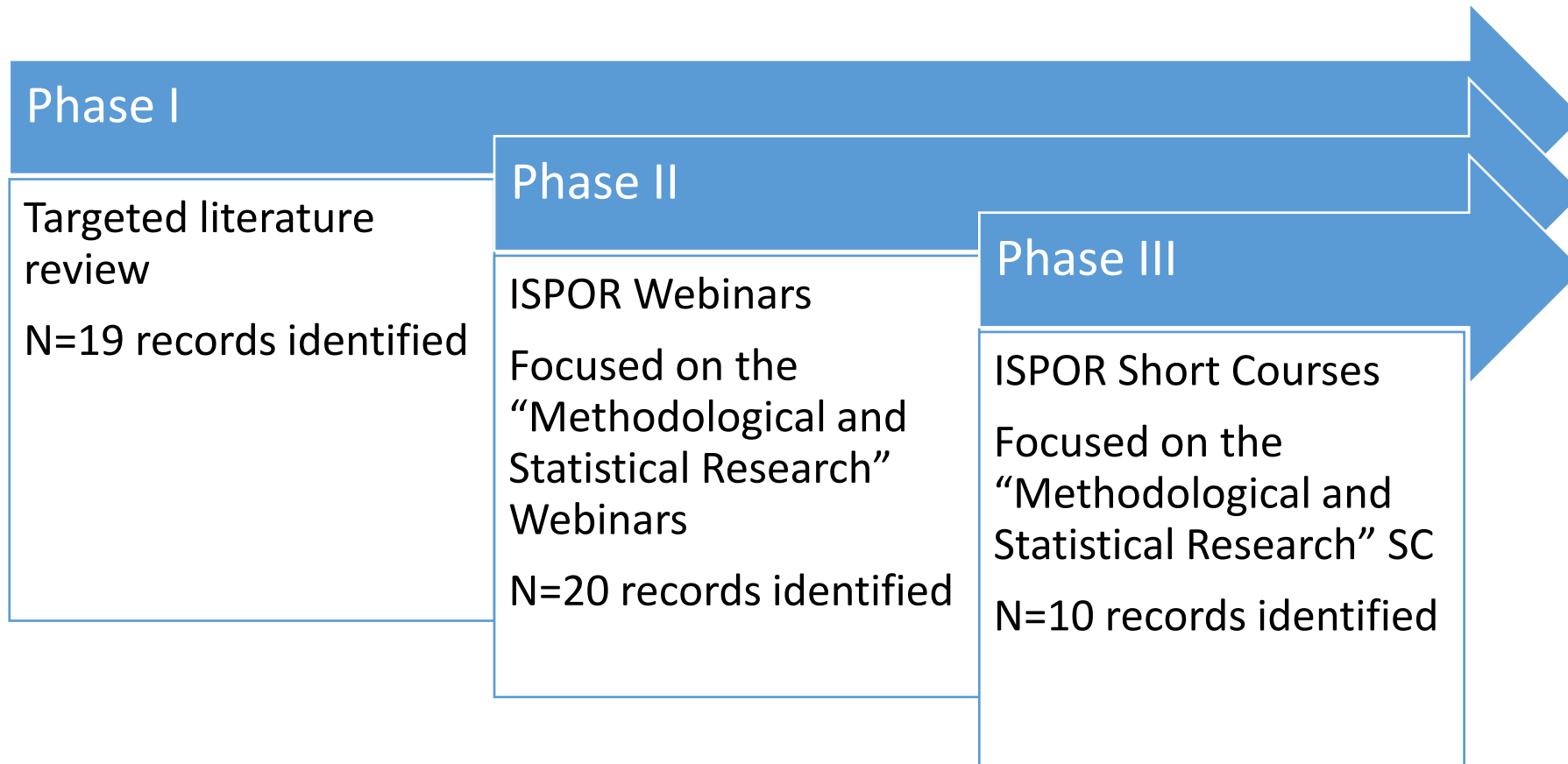
Methodological and Statistical Research

Subtopic: Statistics and Analytics

Statistics and Analytics: Subtopics

- Study design; statistical inference; risk adjustment
- Regression models
- ETL: Extract, Transform, Load
- Data quality
- Big data methods (ML/AI)
- Prediction models
- Statistical programs
- Emerging interfaces
- Communicating results: visualization

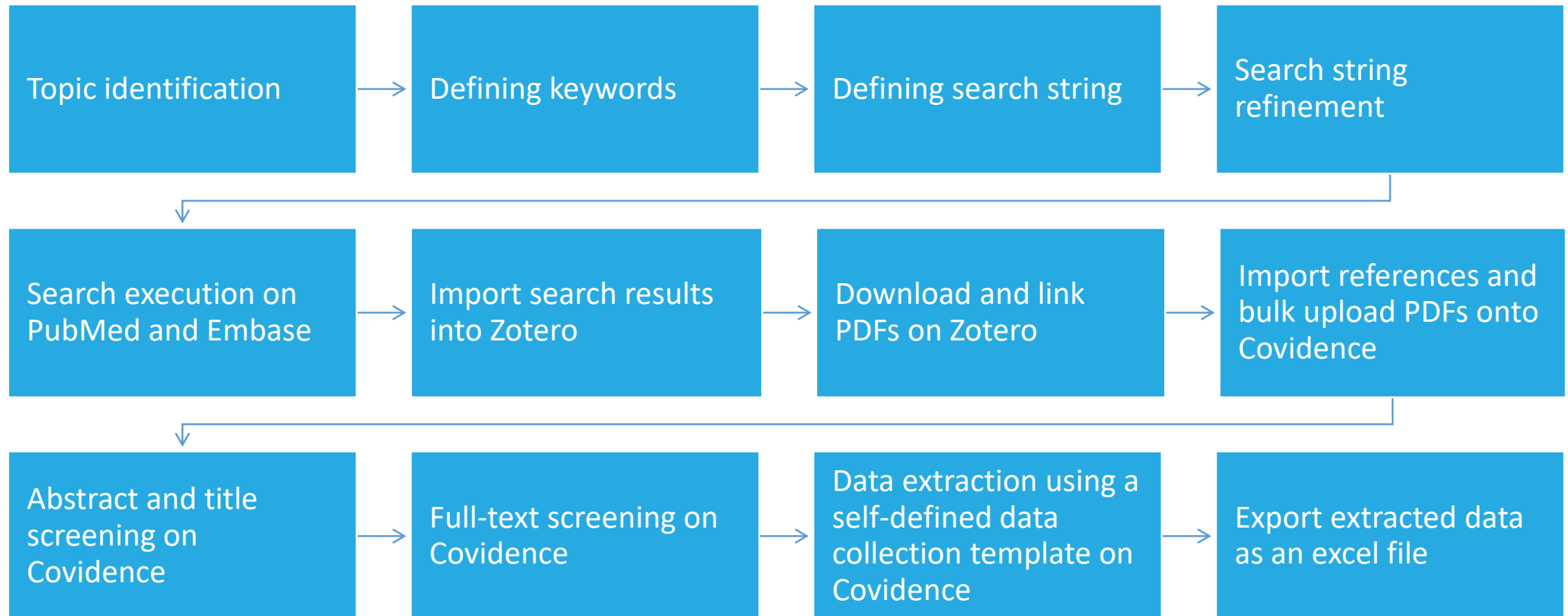
Overview of the Project Phases



2

Literature

Workflow for Published Literature Review



Search Process for Published Literature Review

- A literature review was initiated in PubMed and Embase on Aug 24, 2021
 - Search between 2000-2020
 - Apply criteria of 'Humans' and 'English'
 - Search from selected journals
 - Use the Medical Subject Heading (MeSH) search terms: 'methods', 'statistics as topic/methods', 'statistics as topic', or 'methods/statistics and numerical data'
 - Search titles and abstracts for the combinations of identified topic keywords

Selected Journals for the review

- Value in Health
- Pharmacoeconomics
- American Journal of Epidemiology
- Health Services Research
- Medical Decision Making
- Pharmacoepidemiology and Drug Safety
- Health Services Outcomes Research Methodology
- International Journal of Technology Assessment in Health Care

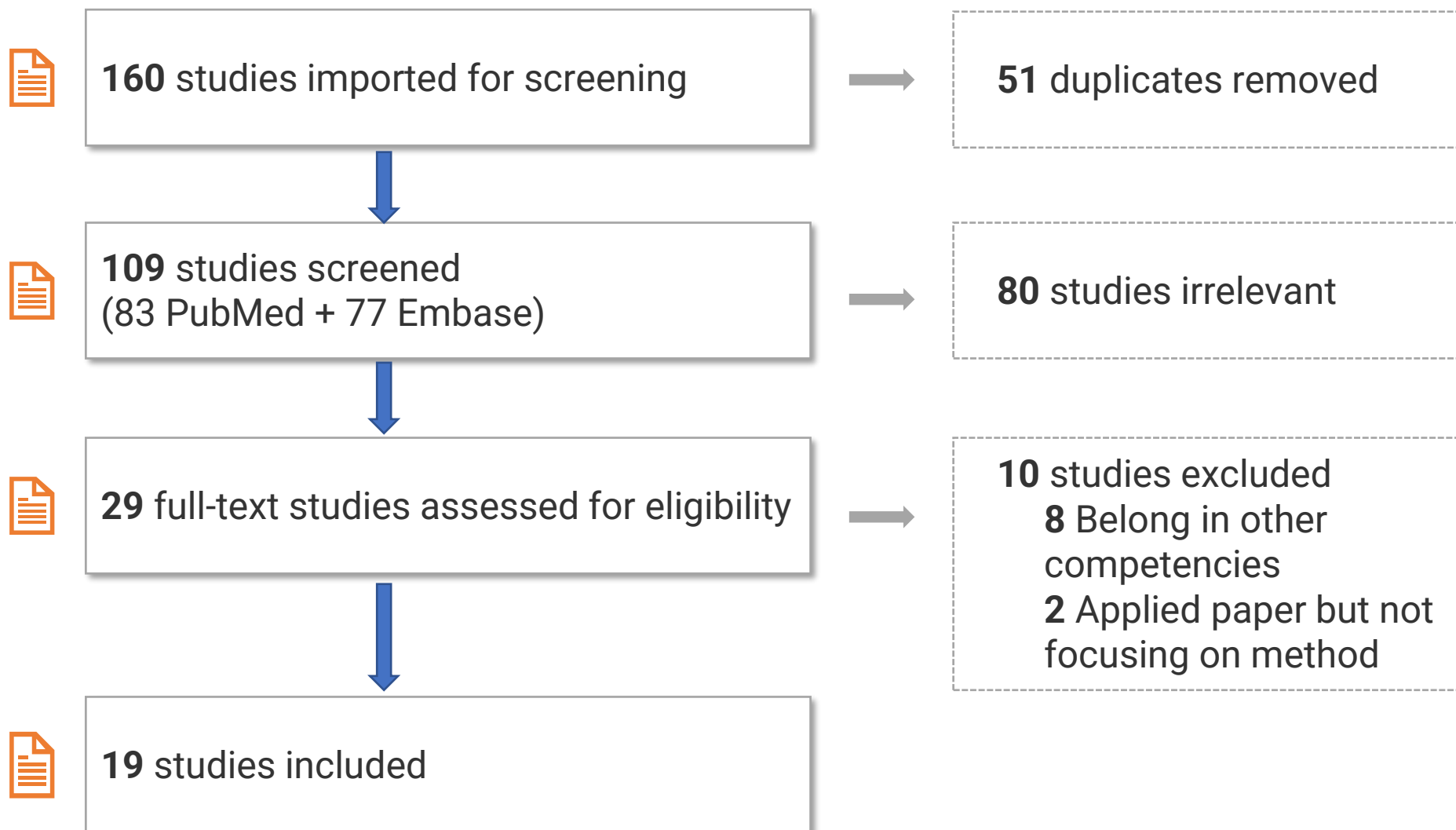
Search Terms for Title and Title/Abstract

Title: Tutorial / Primer / Introduction / Guide / State of the Art

Title/Abstract

- Study design observational research
- Regression
- Statistical inference
- Machine learning
- Artificial intelligence
- Prediction
- Statistical programs
- Risk adjustment
- Survival
- Mixed model
- Marginal structural model
- Count model
- Propensity score
- Disease risk score
- Instrumental variable
- Confounding
- Bias
- Specification
- Repeated measures
- Missing data
- Imputation
- Simulation
- Calibration
- Extrapolation
- Data visualization
- Bayes

Search Results for Published Literature Review



Data Extraction using Covidence

← Latimer 2013

Save Complete ?

Select Full Text ▾

Check for updates

ORIGINAL ARTICLES

Survival Analysis for Economic Evaluations Alongside Clinical Trials—Extrapolation with Patient-Level Data: Inconsistencies, Limitations, and a Practical Guide

Nicholas R. Latimer, MSc

Background. In health technology assessments (HTAs) of interventions that affect survival, it is essential to accurately estimate the survival benefit associated with the new treatment. Generally, trial data must be extrapolated, and many models are available for this purpose. The choice of extrapolation model is critical because different models can lead to very different cost-effectiveness results. A failure to systematically justify the chosen model creates the possibility of bias and inconsistency between HTAs. **Objective.** To demonstrate the limitations and inconsistencies associated with the survival analysis component of HTAs and to propose a process guide that will help exclude these from future analyses. **Methods.** We reviewed the survival analysis component of 45 HTAs undertaken for the National Institute for Health and Clinical Excellence (NICE) in the cancer disease area. We drew upon our findings to identify common limitations and to develop a process guide. **Results.** The chosen survival models were not systematically justified in any of the HTAs reviewed. The range of models considered was usually insufficient, and the rationale for the chosen model was universally limited; in particular, the plausibility of the extrapolated portion of fitted survival curves was very rarely explicitly considered. **Limitations.** We do not seek to describe and review all methods available for performing survival analysis—several approaches exist that are not mentioned in this article. Instead we seek to analyze methods commonly used in HTAs and limitations associated with their application. **Conclusions.** Survival analysis has not been conducted systematically in HTAs. A systematic approach such as the one proposed here is required to reduce the possibility of bias in cost-effectiveness results and inconsistency between technology assessments. **Key words:** oncology; cost-utility analysis; technology assessment; pharmacoeconomics; modeling; survival analysis; overall survival; resource allocation (*Med Decis Making* 2013;33:743–754)

Received 6 March 2012 from Health Economics and Decision Science, School of Health and Related Research, University of Sheffield (NFL). Presentations based upon this work have been given at the International Health Economics Association (IHEA) 8th World Congress in Toronto, Ontario, Canada.

Interventions that affect survival or any other time-to-event outcome form a high proportion of the treatments appraised by health technology assessment (HTA) agencies. In these it is essential to accu-

DATA EXTRACTION

General information

PubMed ID

Abbreviated Reference
Enter in the format of: [Title]. [Journal].

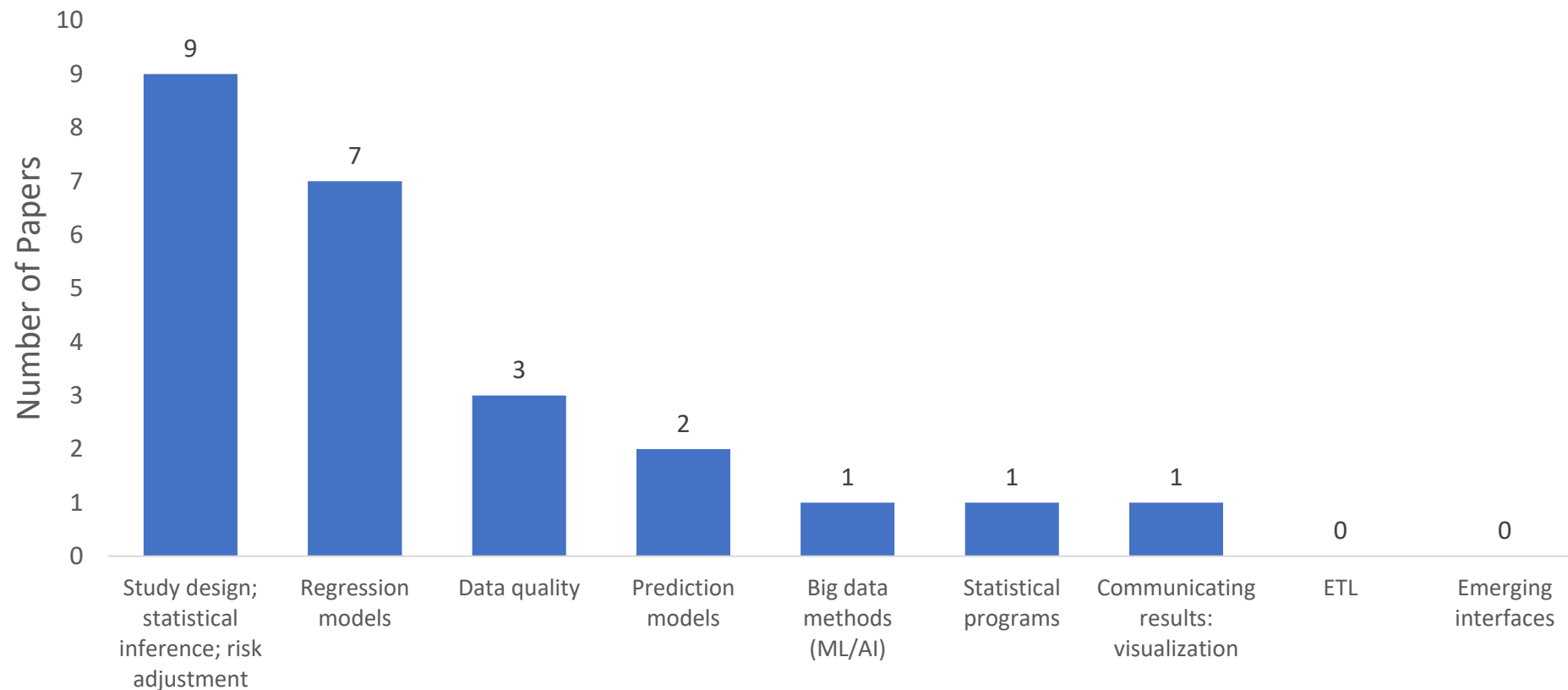
Publication year

Characteristics of included studies

Statistical topic area: ☐ 1 Study design ☐ 2 Regression models ☐ 3 ETL ☐ 4 Data ☐ 5 Statistical inference ☐ 6 Big data methods (ML/AI) ☐ 7 Prediction models ☐ 8 Statistical programs

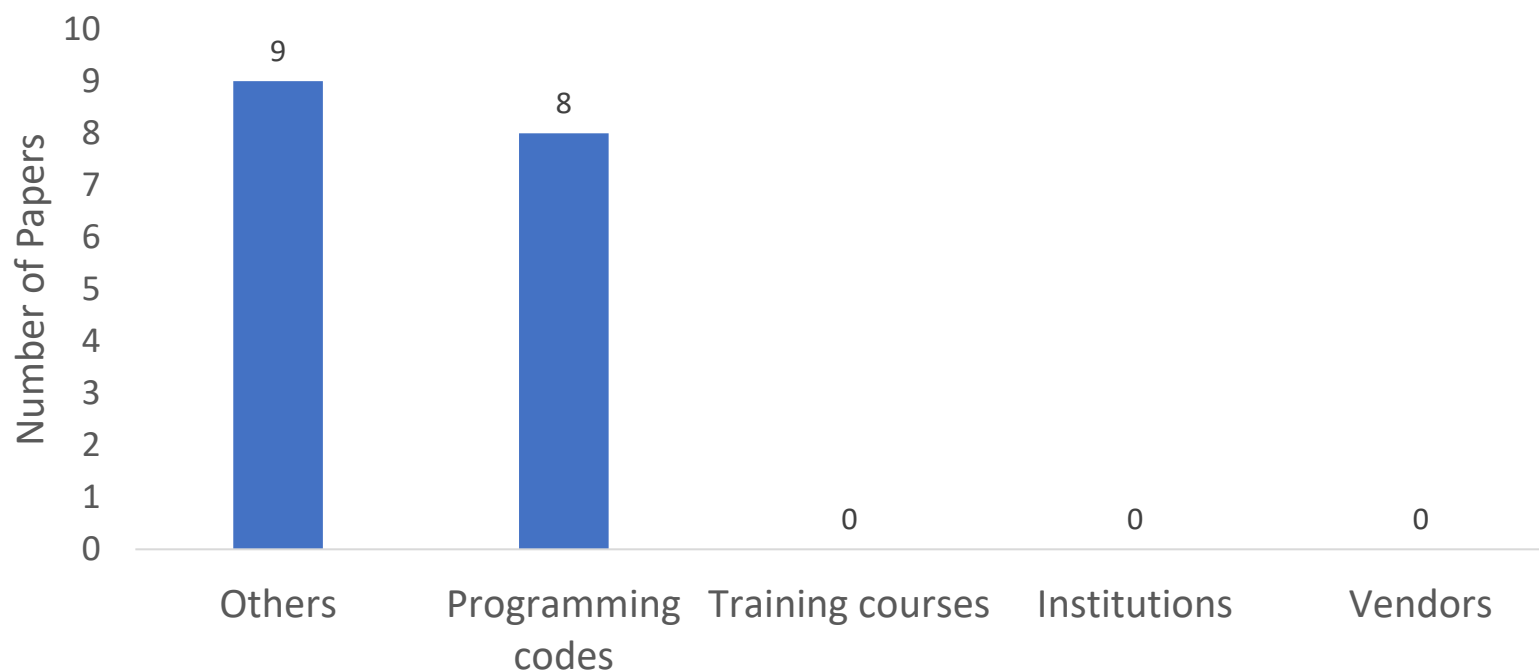
15

Figure 1. Statistical Topic Area Covered (N=19)



- 19 published peer-reviewed papers were evaluated. 5/19 (26%) of papers covered more than one topic area.
- No papers covered ETL or emerging interfaces.

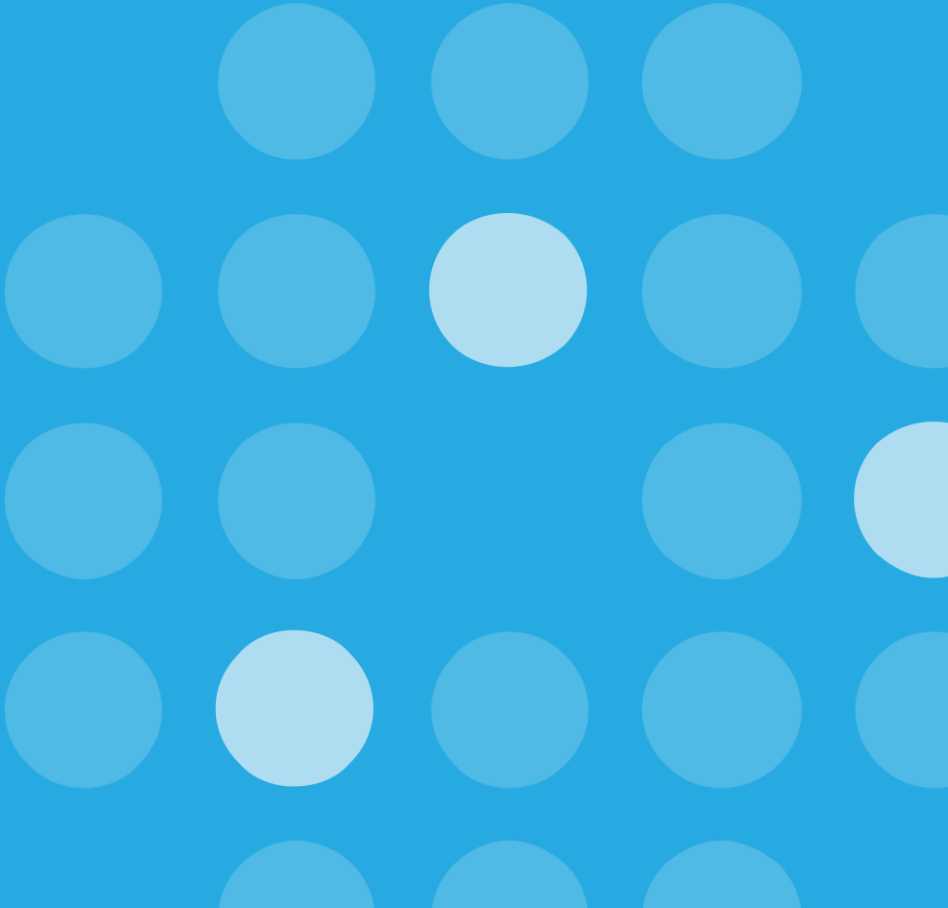
Figure 2. Additional Resources Provided (N=19)



- 5/19 (26%) of papers did not provide additional resources to further participants' learning.
- 3/14 (21%) of papers provided more than one resource.

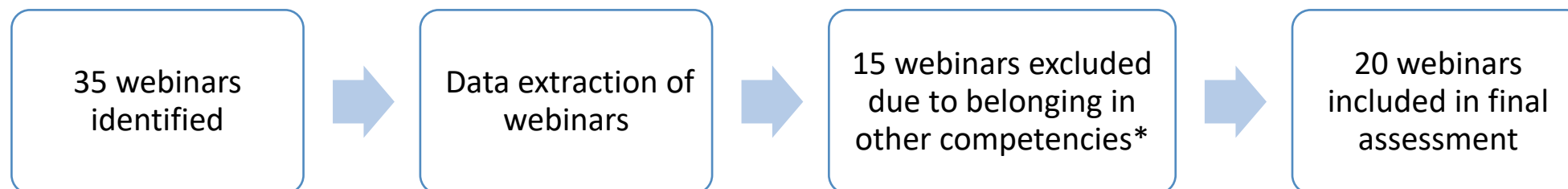
3

Webinars



Search Process for Webinars

- Webinars were identified from the ISPOR website on the webpage of Health Economics and Outcomes Research Webinars
- Eligible criteria
 - Under the section of 'Methodical & Statistical Research'
 - Published between 2012-2020
- Recordings of the webinars were reviewed

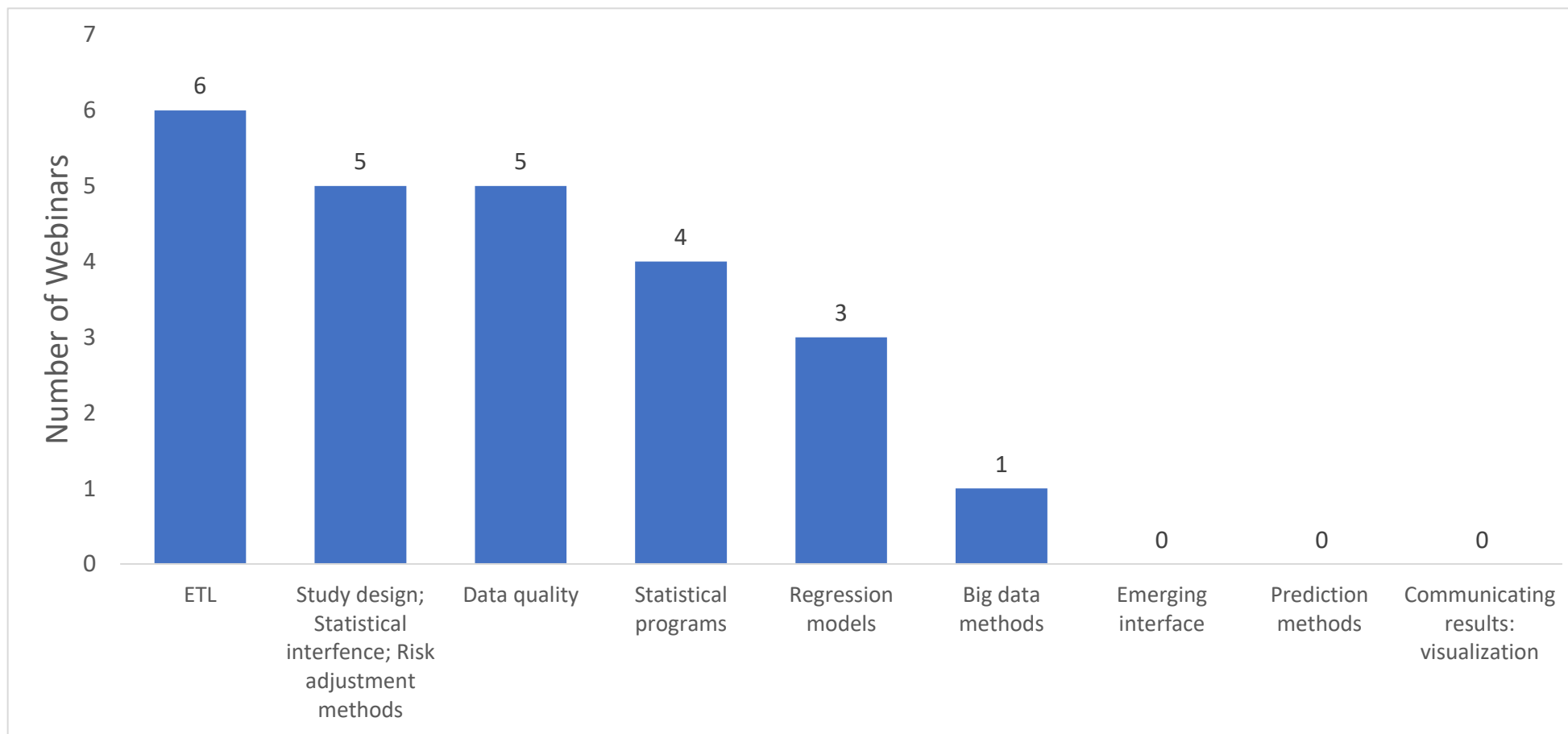


* Other competencies of the excluded webinars include simulation modeling, constrained optimization, decision analysis, qualitative content analysis.

Examples of Webinar Titles

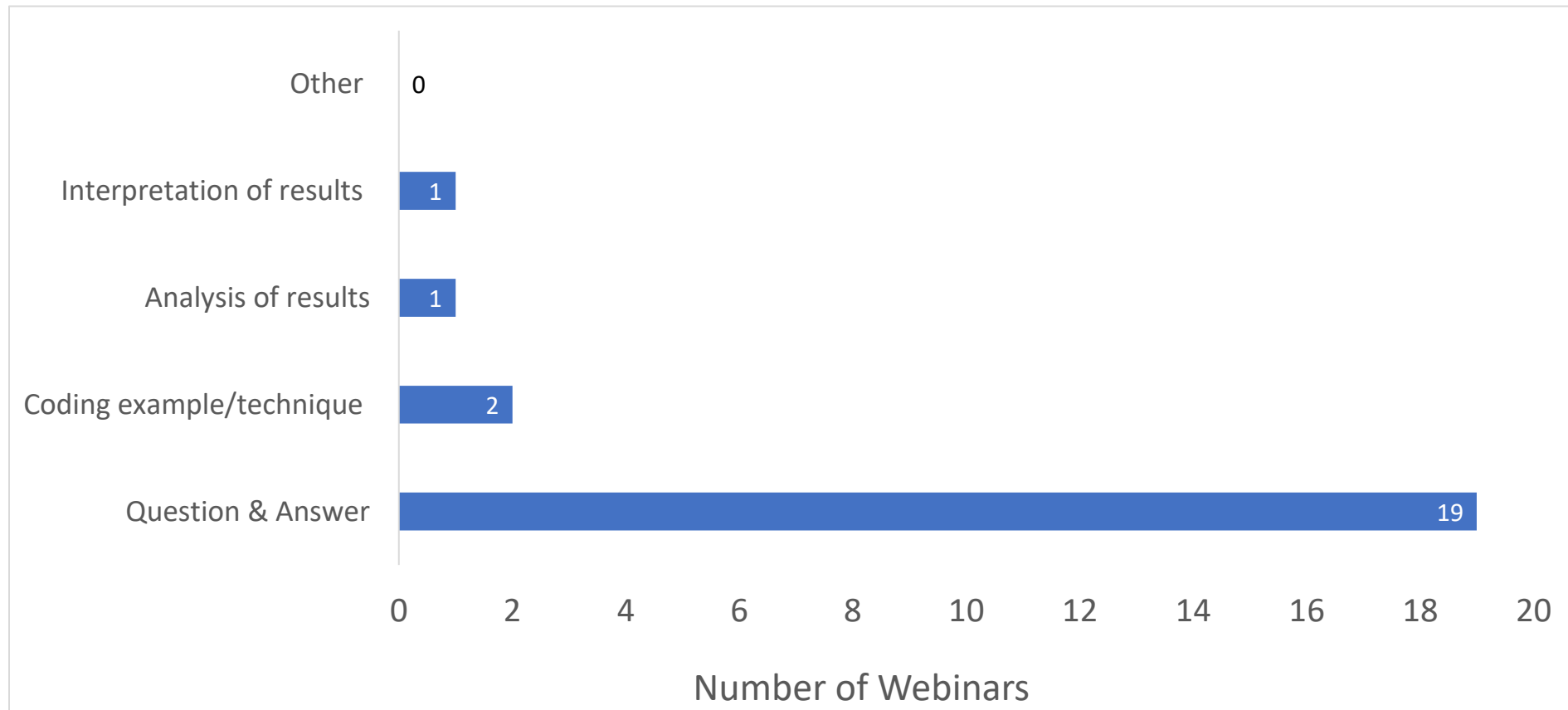
- An Introduction to Cost Analysis Methods Using Large Administrative Claims Databases
- Consideration in Deriving Propensity Score for Comparative Effectiveness or Safety Studies
- Machine Learning in HEOR
- Extrapolation in Survival Analysis: What? Why? How? Where next?
- Instrumental Variables Estimation in Pharmacoeconomics and Outcomes Research

Figure 3. Statistical Topic Area Covered (N=20)



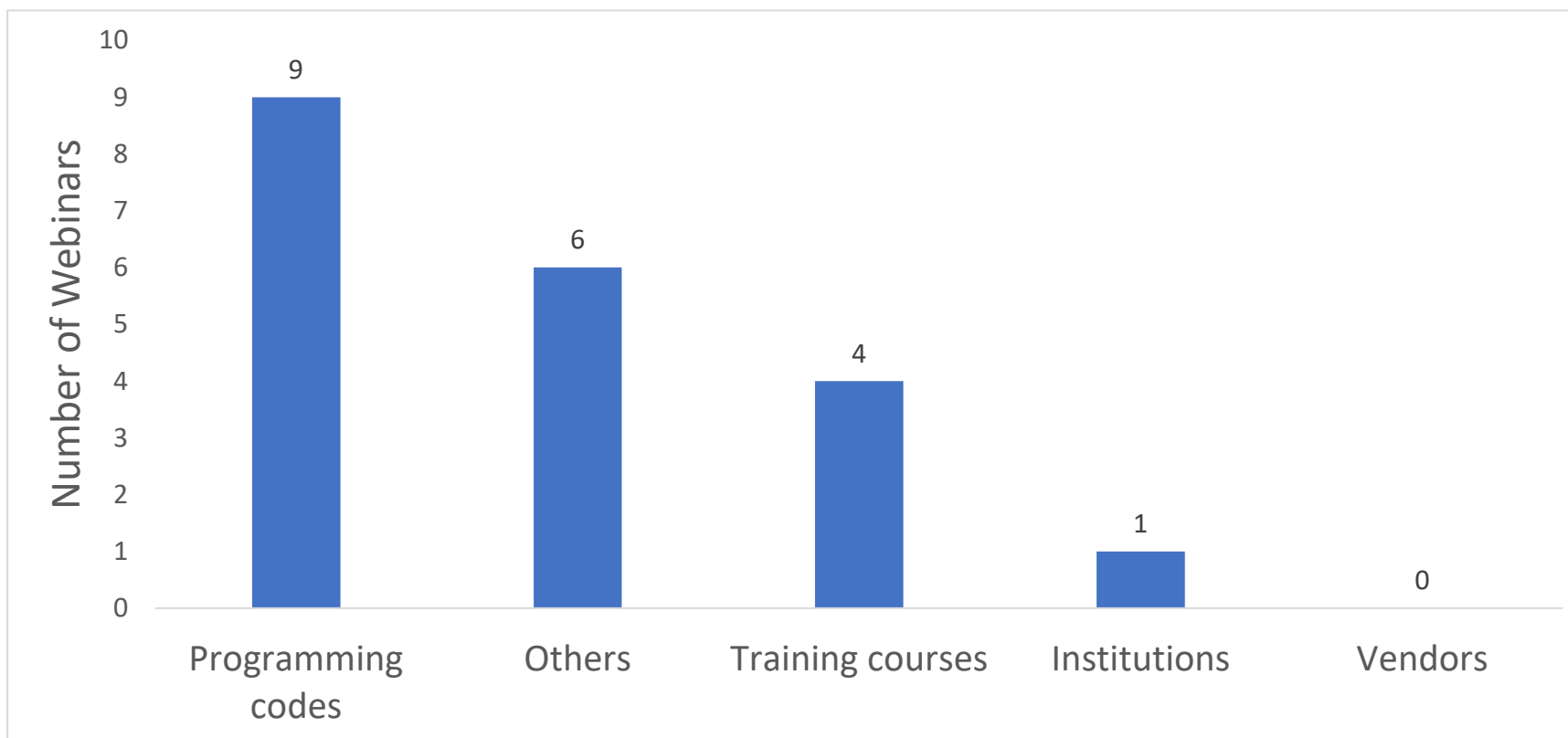
- 20 webinars were evaluated, 5/20 (25%) of webinars covered more than one topic area.
- No webinars covered emerging interfaces, prediction methods, communication results (visualization).

Figure 4. Type of Interaction (N=20)



- All 20 webinars included a form of interaction.
- 3/20 (15%) of webinars included more than one form of interaction.

Figure 5. Additional Resources Provided (N=20)



- 9/20 (45%) of webinars did not provide any additional resources to further participants learning after the webinar.
- 7/11 (63%) of webinars provided more than one resource.

SECTION

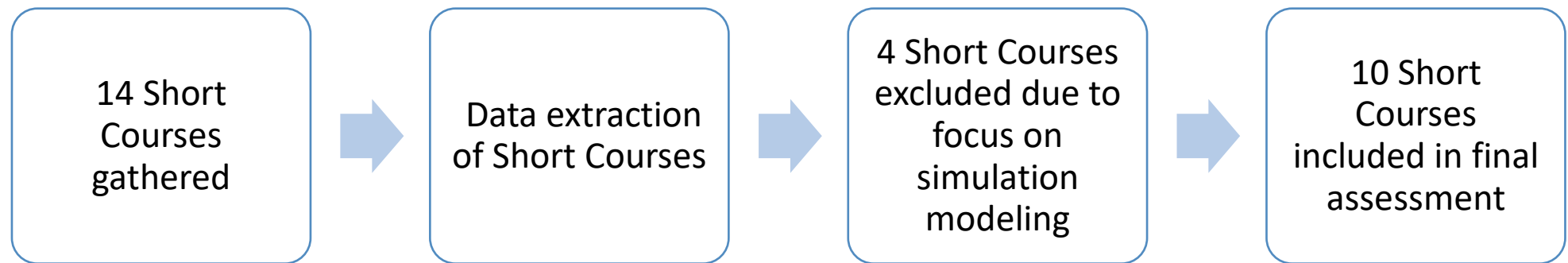
4

Short Courses



Selection of Short Courses

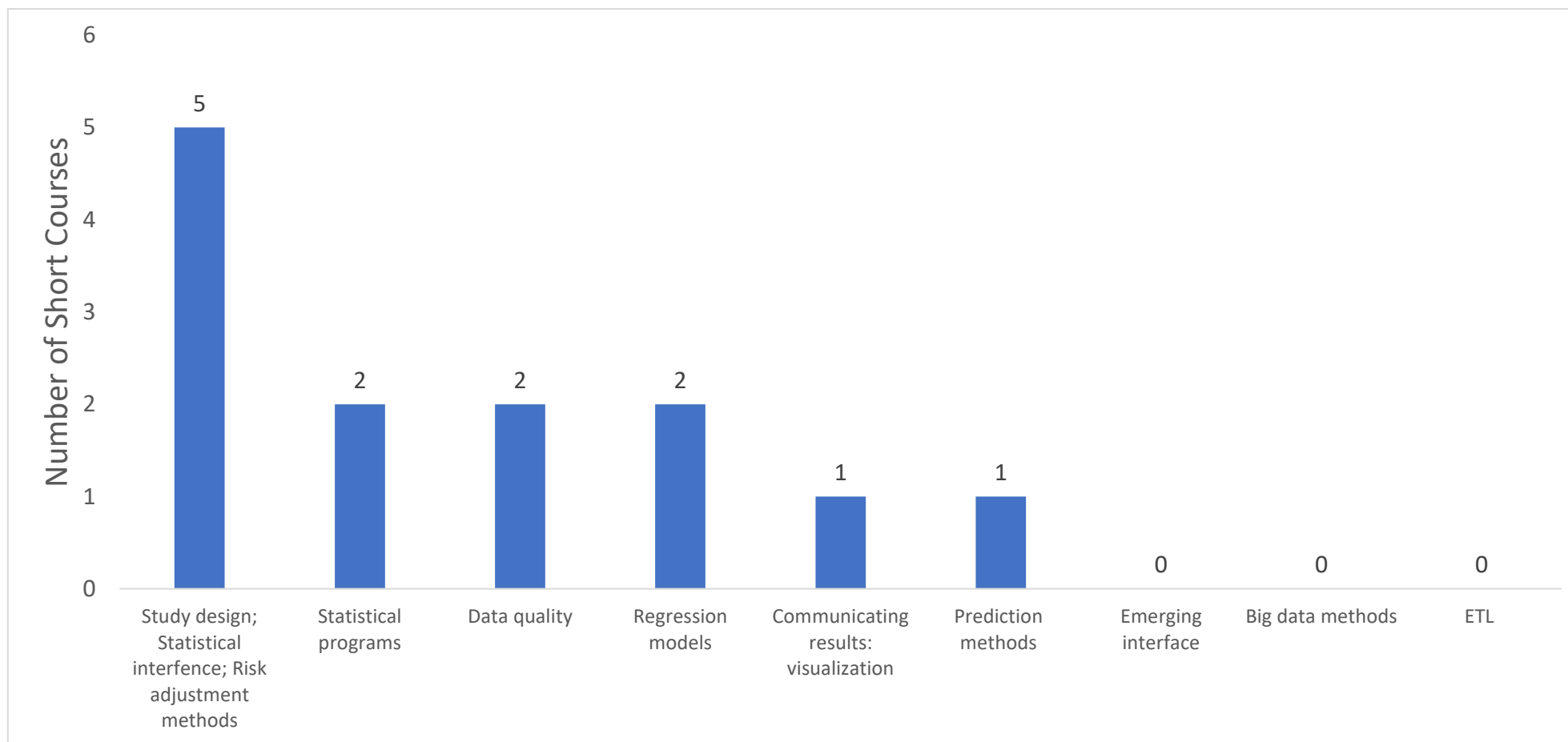
- ISPOR's internal short course database used as the sole source for short courses
- Courses within the Methodological & Statistical Research track were included
- Short Courses between 2018-2021 were included for data extraction
- PDF versions of short courses were used for data extraction



Short Course Titles – Full list

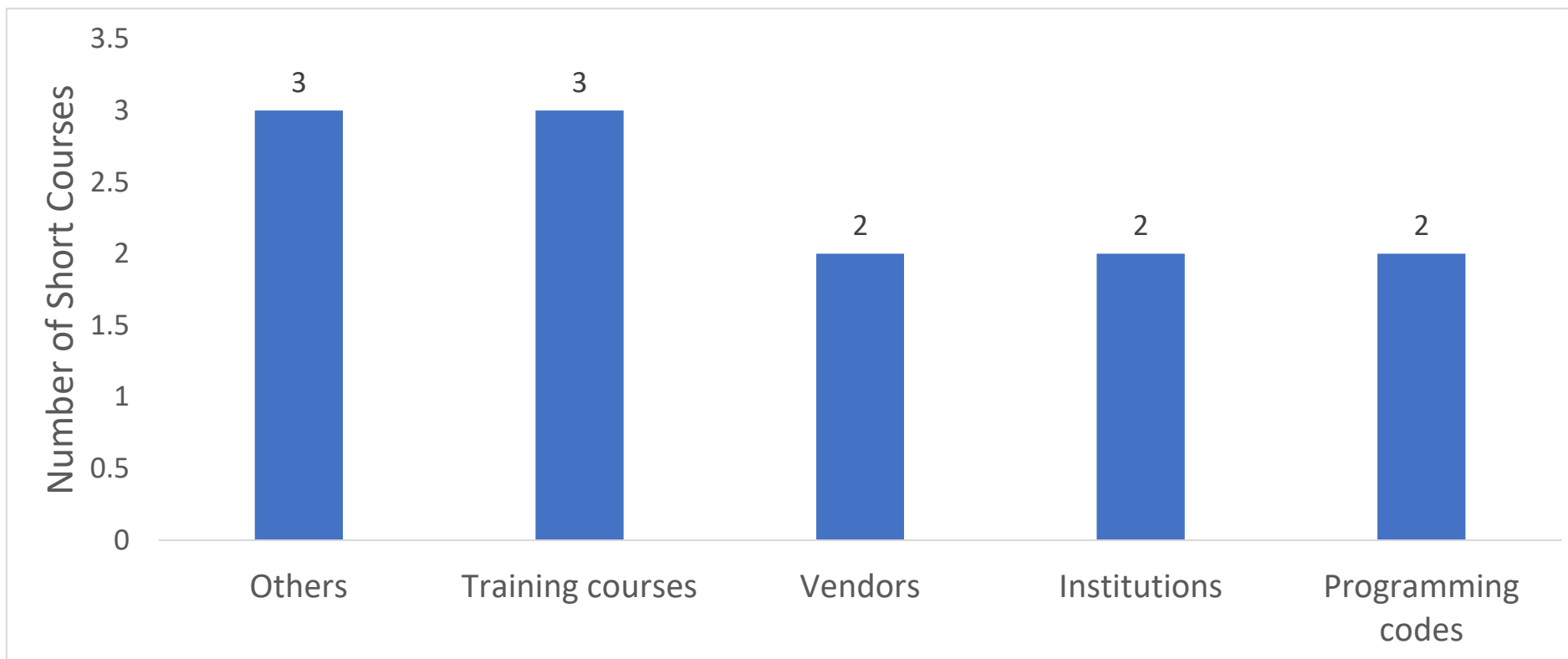
Title	Level
Development of Conceptual Models	Introductory
Introduction to Constrained Optimization Methods for Health Care Research	Introductory
Introduction to Modeling	Introductory
Bayesian Analysis – Overview and Applications	Introductory
Bayesian Analysis - Overview and Applications	Intermediate
Adjusting for Time-Dependent Confounding and Treatment Switching Bias in Observational Studies and Clinical Trials: Purpose, Methods, Good Practices and Acceptance in HTA	Intermediate
Advanced Methods for Addressing Selection Bias in Real-World Effectiveness and Cost-Effectiveness Studies	Intermediate
Analysis of Longitudinal Data: Fixed and Random Effects Models	Intermediate
Understanding Survival Modeling with Application to HTA	Intermediate
Advanced Decision Modeling for Health Economic Evaluations	Advanced

Figure 6. Statistical Topic Area Covered (N=10)



- 10 short courses were evaluated, 3/10 (30%) of short courses covered more than on statistical topic area.
- No short courses covered emerging interfaces, big data methods, or ETL

Figure 7. Additional Resources Provided (N=10)

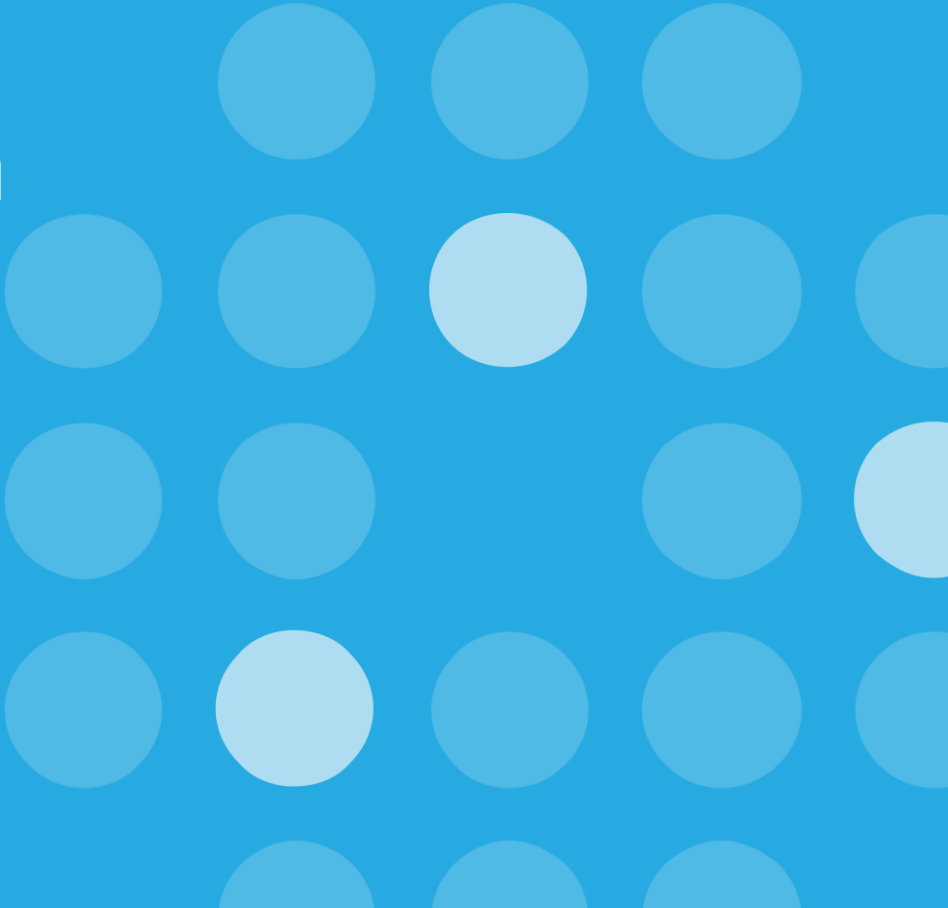


- 4/10 (40%) of short courses did not provide additional resources to further participants' learning after the short course.
- 4/6 (66%) provided more than one additional resource.

SECTION

5

Discussion



Audience discussion: Training gaps and opportunities

Topic	Literature	Webinars	Short Courses
Study design; statistical inference; risk adjustment	9	5	5
Regression models	7	3	2
ETL: Extract, Transform, Load	0	6	0
Data quality	3	5	2
Big data methods (ML/AI)	1	1	0
Prediction models	2	0	1
Statistical programs	1	4	2
Emerging interfaces	0	0	0
Communicating results: visualization	1	0	1

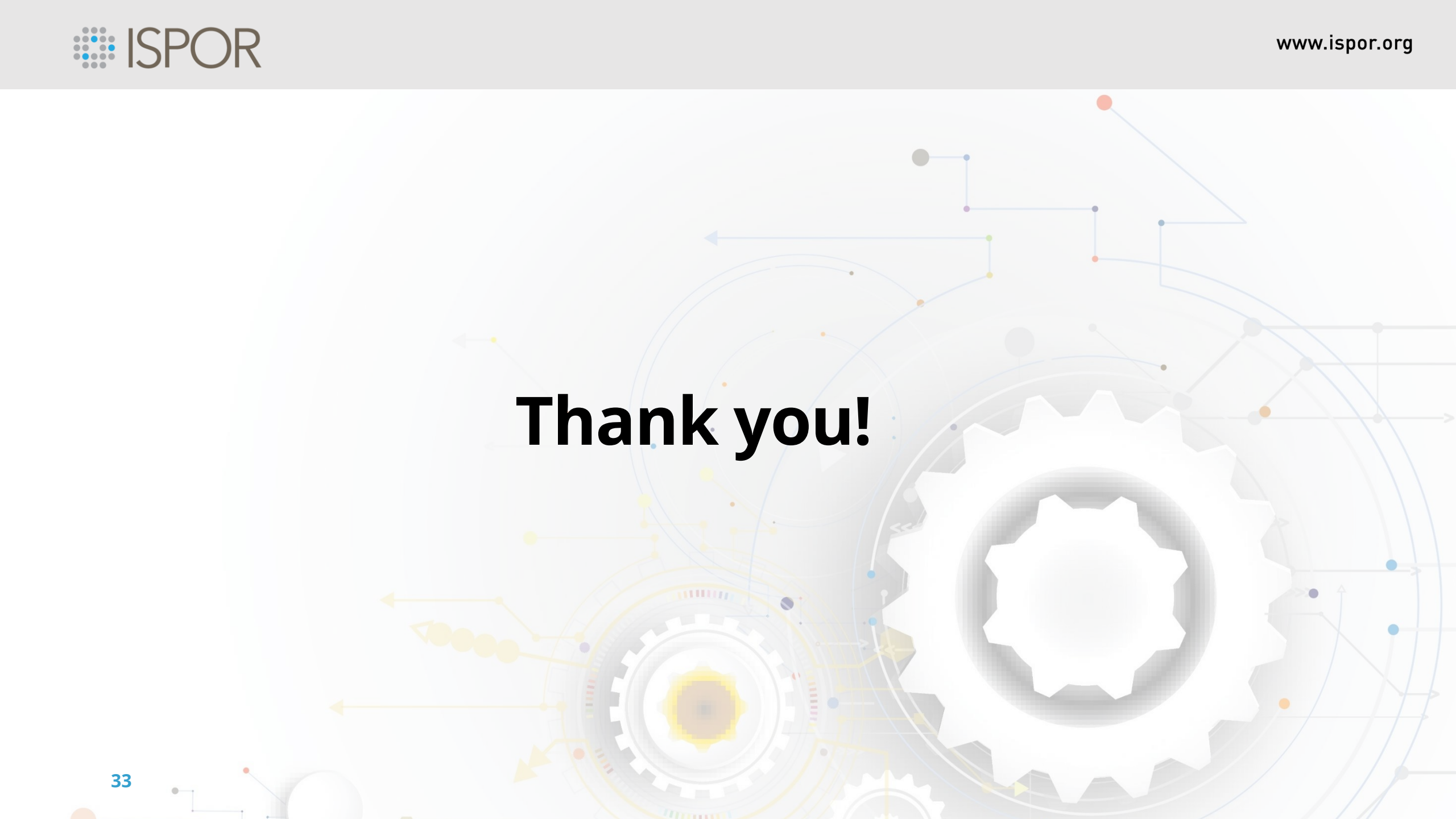
Audience discussion – Training resources

- What types of 'Statistics and Analytics' training resources **have been helpful to you in your work?**
- In your current position, what types of 'Statistics and Analytics' training resources **would be helpful to you in your work?**

Audience discussion: Next category

- What would be the next competency category to address?

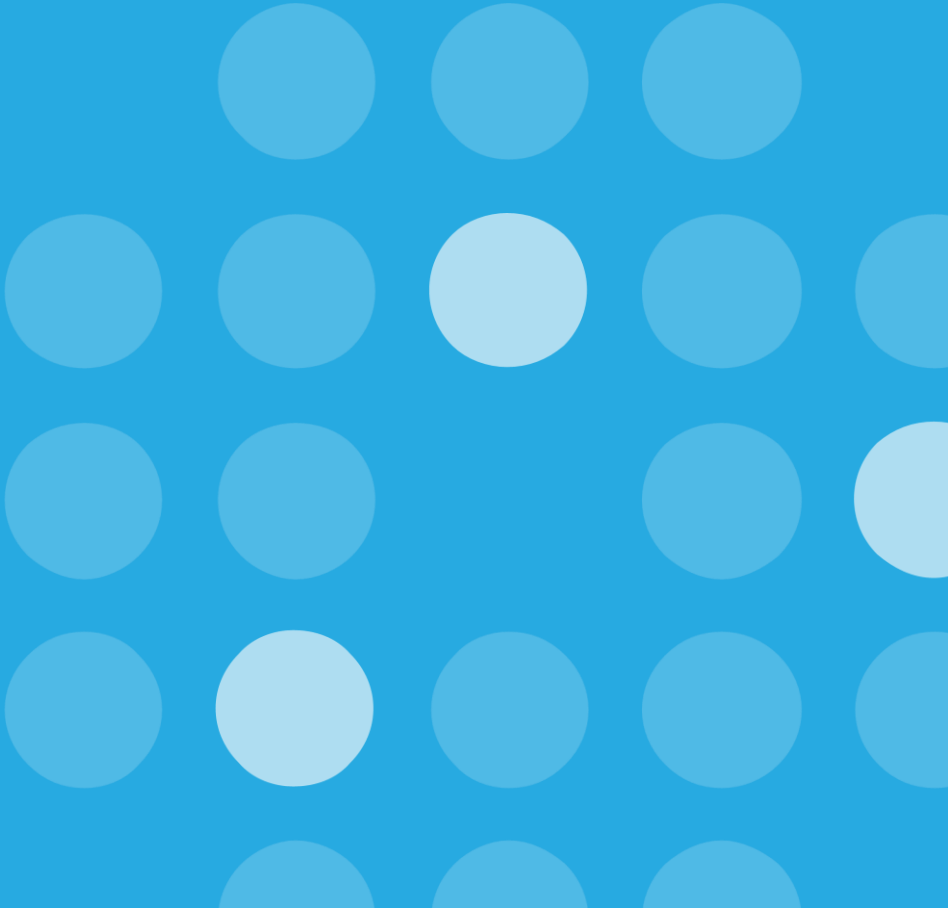




Thank you!

6

Appendix

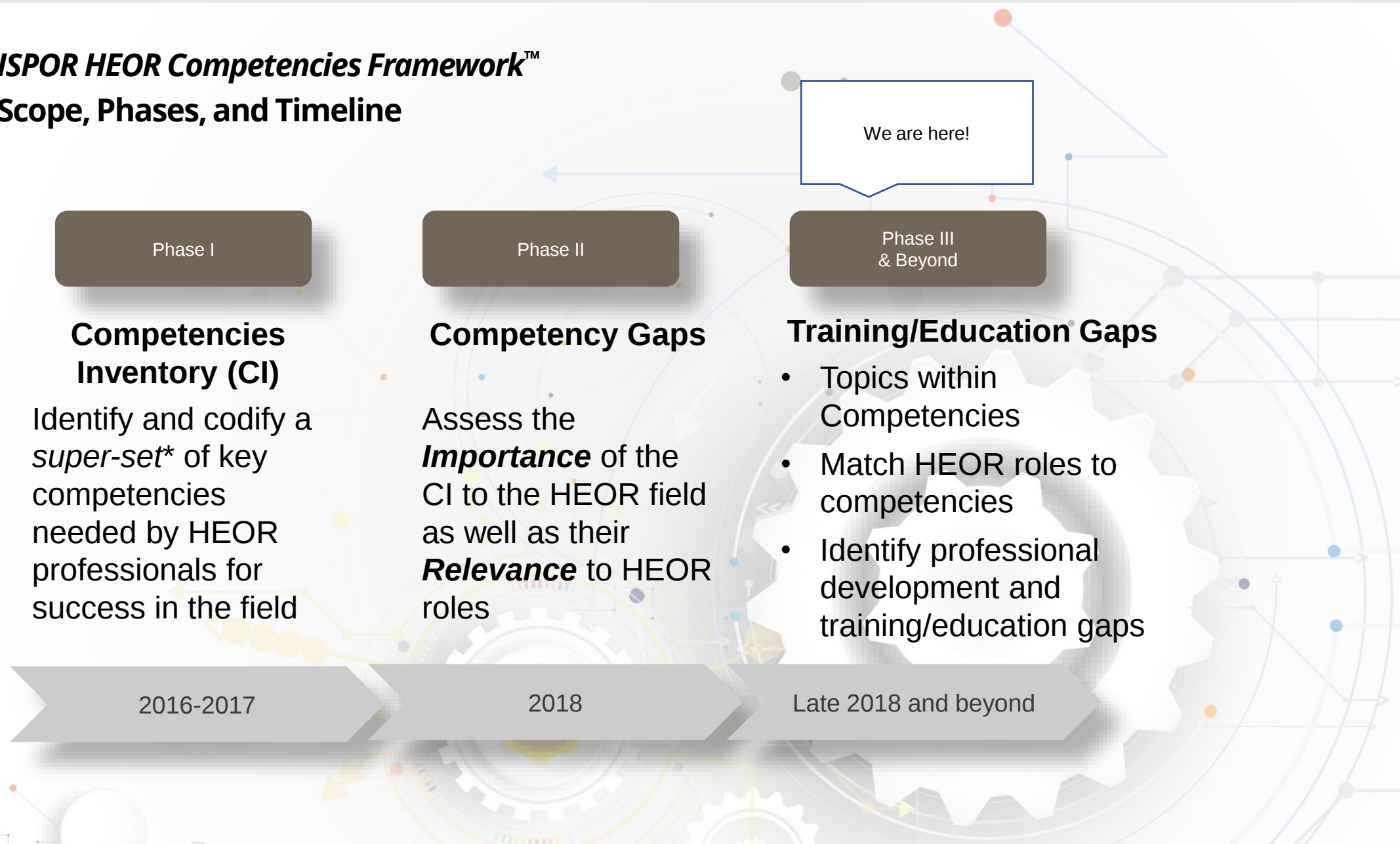


Relevance of the Competency Framework to ISPOR's Mission

- Mission: To promote health economics and outcomes **research excellence** to **improve decision making** for health globally
 - Competency Framework promotes **research excellence** by outlining the knowledge and skills necessary to conduct and report health economics and outcomes research
 - Competency Framework promotes **improved decision making** by specifying that an understanding of the healthcare system, health technology assessment, and HEOR communication are necessary for professionals in our field



ISPOR HEOR Competencies Framework™ Scope, Phases, and Timeline



Appendix. Search strings used (PubMed)

2000/01/01:2020/12/31[Date - Publication] **AND** "english"[Language] **AND** "humans"[Filter] **AND** ("Value in Health the journal of the international society for pharmacoeconomics and outcomes research"[Journal] OR "PharmacoEconomics"[Journal] OR "American Journal of Epidemiology"[Journal] OR "Health Services Research"[Journal] OR "Medical Decision Making an international journal of the society for medical decision making"[Journal] OR "Pharmacoepidemiology and Drug Safety"[Journal] OR "Health Services Outcomes Research Methodology"[Journal] OR "Int J Technol Assess Health Care"[Journal])

AND ("tutorial"[Title] OR "primer"[Title] OR "introduction"[Title] OR "guide"[Title] OR "state of the art"[Title])

AND ("study design observational research"[Title/Abstract] OR "regression"[Title/Abstract] OR "statistical inference"[Title/Abstract] OR "machine learning"[Title/Abstract] OR "artificial intelligence"[Title/Abstract] OR "prediction"[Title/Abstract] OR "statistical programs"[Title/Abstract] OR "risk adjustment"[Title/Abstract] OR "survival"[Title/Abstract] OR "mixed model"[Title/Abstract] OR "marginal structural model"[Title/Abstract] OR "count model"[Title/Abstract] OR "propensity score"[Title/Abstract] OR "disease risk score"[Title/Abstract] OR "instrumental variable"[Title/Abstract] OR "confounding"[Title/Abstract] OR "bias"[Title/Abstract] OR "specification"[Title/Abstract] OR "repeated measures"[Title/Abstract] OR "missing data"[Title/Abstract] OR "imputation"[Title/Abstract] OR "simulation"[Title/Abstract] OR "calibration"[Title/Abstract] OR "extrapolation"[Title/Abstract] OR "data visualization"[Title/Abstract] OR "Bayes"[Title/Abstract] OR "methods"[MeSH Terms] OR "statistics as topic/methods"[MeSH Terms] OR "statistics as topic"[MeSH Terms] OR "methods/statistics and numerical data"[MeSH Terms])

Appendix. Search strings used (Embase)

[2000-2020]/py **AND** [english]/lim **AND** [humans]/lim **AND** ([article]/lim OR [article in press]/lim OR [editorial]/lim OR [review]/lim) **AND** ('Value in Health'/jt OR PharmacoEconomics/jt OR 'American Journal of Epidemiology'/jt OR 'Health Services Research'/jt OR 'Medical Decision Making : an international journal of the society for medical decision making'/jt OR 'Pharmacoepidemiology and Drug Safety'/jt OR 'Health Services and Outcomes Research Methodology'/jt OR 'International Journal of Technology Assessment in Health Care'/jt) **AND** ('tutorial':ti OR 'primer':ti OR 'introduction':ti OR 'guide':ti OR 'state of the art':ti) **AND** ('study design observational research':ti,ab OR 'regression':ti,ab OR 'statistical inference':ti,ab OR 'machine learning':ti,ab OR 'artificial intelligence':ti,ab OR 'prediction':ti,ab OR 'statistical programs':ti,ab OR 'risk adjustment':ti,ab OR 'survival':ti,ab OR 'mixed model':ti,ab OR 'marginal structural model':ti,ab OR 'count model':ti,ab OR 'propensity score':ti,ab OR 'disease risk score':ti,ab OR 'instrumental variable':ti,ab OR 'confounding':ti,ab OR 'bias':ti,ab OR 'specification':ti,ab OR 'repeated measures':ti,ab OR 'missing data':ti,ab OR 'imputation':ti,ab OR 'simulation':ti,ab OR 'calibration':ti,ab OR 'extrapolation':ti,ab OR 'data visualization':ti,ab OR 'Bayes':ti,ab OR 'methodology'/exp OR 'methodological studies'/exp OR 'statistical concepts'/exp OR 'statistics'/exp)

Statistical Topics	Description
Study design	"The "architecture," "anatomy," or "physiology" of a study: its structure and procedures, the specific methods, details on the selection and management of the study population (e.g., follow-up), time frame, and ethical decisions." (Ref: Porta, M. (2014). A Dictionary of Epidemiology.)
Regression models	Regression analysis is a statistical tool for evaluating the relationship of one or more independent variables to a single dependent variable. (Ref: Kleinbaum, D. G., Kupper, L. L., Nizam, A., & Rosenberg, E. S. (2014). Applied regression analysis and other multivariable methods.)
Extract, transform, and load (ETL)	General framework to extract data from various sources, transform the data using standardized methods, and then store the data in an appropriate setting
Data	Data - "A collection of items of information." Data cleaning - "The process of excluding the information in incomplete, inconsistent records or irrelevant information collected in a survey or other form of epidemiological study before analysis begins." Data processing - "Conversion of items of information into a form that permits storage, retrieval, and analysis." (Ref: Porta, M. (2014). A Dictionary of Epidemiology.)
Statistical inference	"Statistical inference is to draw a conclusion about a population of items or measurements based on information contained in a sample from that population." "Statistical inferences are made through estimation and hypothesis testing." (Ref: Kleinbaum, D. G., Kupper, L. L., Nizam, A., & Rosenberg, E. S. (2014). Applied regression analysis and other multivariable methods.)
Big data methods (ML/AI)	Machine learning - "The ability of a program to learn from experience—i.e., to modify its execution on the basis of newly acquired information. In epidemiology and bioinformatics, examples include artificial neural networks, support vector machines, Bayesian networks, and other methods that update their procedures as new data are provided." Artificial intelligence - "A branch of computer science in which attempts are made to duplicate human intellectual functions. One application is in clinical diagnosis, in which computer programs are based upon statistical analyses of data abstracted from clinical records." (Ref: Porta, M. (2014). A Dictionary of Epidemiology.)
Prediction models	Predictive modeling - "the process of applying a statistical model or data mining algorithm to data for the purpose of predicting new or future observations." "the goal is to predict the output value (Y) for new observations given their input values (X)" "Predictions include point or interval predictions, prediction regions, predictive distributions, or rankings of new observations. Predictive model is any method that produces predictions, regardless of its underlying approach: Bayesian or frequentist, parametric or nonparametric, data mining algorithm or statistical model, etc." (Ref: Shmueli, G. (2010). To explain or to predict? SSRN Electronic Journal.)
Statistical programs	Computer programs or softwares that can be used for statistical analyses of data.
Emerging interfaces	Novel interfaces between a human user and a statistical program or software to perform statistical procedures.
Risk adjustment methods	"A method of accounting for the morbidity of individual patients when making aggregate comparisons between population outcomes. Risk adjustment uses complex algorithm and models to remove or reduce the effects of a confounding factor or risk factor, when assessing outcomes." "Varying types of outcomes that ranges from resource use such as cost and length of stay to mortality, may be adjusted for risk." (Ref: ISPOR BOOK OF TERMS 2004)
Communicating results (visualization)	The schematic representation of analytical results through infographics such as charts or diagrams that can be used to communicate information better and faster than written text.