The terminology of clinical reasoning in health professions education: implications and considerations

Abstract:

Introduction: Clinical reasoning is considered to be at the core of health practice. Here, we report on the diversity and inferred meanings of the terms used to refer to clinical reasoning, and consider implications for teaching and assessment.

Methods: In the context of a Best Evidence Medical Education (BEME) review of 625 papers drawn from 18 health professions, we identified 110 terms for clinical reasoning. We focus on iterative categorization of these terms across three phases of coding and considerations for how terminology influences educational practices.

Results: Following iterative coding with 5 team members, consensus was possible for 74, majority coding was possible for 16, and full team disagreement existed for 20 terms. Categories of terms included: purpose/goal of reasoning, outcome of reasoning, reasoning performance, reasoning processes, reasoning skills, and context of reasoning.

Discussion: Findings suggest that terms used in reference to clinical reasoning are non-synonymous, not uniformly understood, and the level of agreement differed across terms. If the language we use to describe, to teach, or to assess clinical reasoning is not similarly understood across clinical teachers, program directors, and learners, this could lead to confusion regarding what the educational or assessment targets are for 'clinical reasoning'.

Keywords: Clinical reasoning, assessment, teaching, review, Best Evidence Medical Education

The terminology of clinical reasoning in health professions education research: implications and considerations

Clinical reasoning is considered a core component of being a health professional (Higgs et al 2008), and has been an important focus for teaching and assessment in health professions education (summaries in: Eva 2005; Young et al 2018a). Clinical reasoning has been included within formal competency frameworks across the health professions (e.g. Accreditation Council for Graduate Medical Education Core Competencies, the Royal College of Physicians and Surgeons of Canada's CanMEDS framework, the General Medical Council's Good Medical Practice, the Canadian Association of Occupational Therapists Profile of Practice, the Canadian Physiotherapy Association Competency Profile), underpinning its key role in professional practice (Higgs et al 2008). However, clinical reasoning is not a homogeneous construct - the literature on clinical reasoning is broad and diverse, representing many different professions and theoretical frameworks (Eva 2005; Norman 2005; Blumenthal-Barby et al 2015; Holmboe & Durning 2014; Cutrer et al 2013; Norman et al 2013; Ilgen et al 2012). In addition to several different theoretical frameworks within the literature, there are a large number of terms used to refer to clinical reasoning. A recent review identified over one hundred terms used in reference to clinical reasoning across the health professions (Young et al [under review]). The diversity of terms (e.g. diagnostic reasoning, intuitive processes, contextualized reasoning, critical thinking, shared decision making) leads to different targets for specific teaching interventions, different areas of focus for assessment, and likely different understandings of what clinical reasoning is. These terms are seldom explicitly defined (Young et al [under review]), leaving interpretation up to the reader. Different interpretations of the same term could result in important differences in how trainees and teachers collectively understand clinical reasoning, with resulting differences in the focus of

teaching and assessment. In turn, this raises concerns about comparability of educational experiences, or fairness in assessment.

It is imperative that we have a clear understanding of the construct of clinical reasoning to ensure defensible and sound pedagogical and assessment practices. In order to explore potentially different meanings of terms used to refer to clinical reasoning, we engaged in an iterative coding process to try to categorize terms used to refer to clinical reasoning and refine our understanding of how clinical reasoning is labelled in the Health Professions Education (HPE) literature. The purpose of the current paper is to: 1) present and discuss the variety of terms used in reference to clinical reasoning; 2) report the process of categorization of terms according to inferred meaning; and 3) describe areas of disagreement in the inferred meanings of different terms to better support the teaching and assessment of clinical reasoning.

Methodology

The present study is an in-depth, secondary analysis of data collected for a larger BEME scoping review on clinical reasoning across health professions (details reported in Young et al, under review). This scoping study identified and categorized terms used to refer to clinical reasoning across 625 papers included in the review.

Data extraction:

For each paper included in the larger review (Young et al [under review]), reviewers (n = 13) were asked whether a term other than "clinical reasoning" was used in the paper; if so, to list the term(s). Up to three terms could be reported per paper. The identified terms form the foundation of the data reported here.

Analysis:

Recognizing that the categorization of terminology is likely influenced by particular professional backgrounds, the categorization process (summarized in Figure 1) involved five team members (MY, AT, SL, DG, LG) with various areas of expertise (researcher or clinical educator; physician or occupational therapist; clinical reasoning as a core focus area of research and/or educational responsibility), in order to increase the rigour and trustworthiness of the analytical process.

[Insert Figure 1]

Phase 1: Developing category structure:

1. An initial inductive grouping of the 110 terms into categories was suggested by MY (researcher with a focus on clinical reasoning) using an iterative and inductive coding approach, informed by her knowledge of the clinical reasoning literature.

2. The initial category structure was critically reviewed and refined by AT (occupational therapist, teacher of clinical reasoning, and education researcher) through in person discussions with MY.

3. This adapted category structure was reviewed critically and further refined by SL (clinical educator, teacher of clinical reasoning, and scholar) through in person discussion with MY and AT. Following the completion of this step, we reached consensus on six main categories of terms.

Phase 2: Testing category structure

4. A list of the 110 terms was provided to DG (physician, teacher of clinical reasoning) and LG (researcher with a focus on clinical reasoning) and they were asked to independently assign each term to one of the six categories.

5. Terms and their category assignments were consolidated across three team members (DG, LG, MY), with focus on which terms were unanimously assigned to the same categories (i.e. team consensus). Terms where consensus coding was present were identified (N=43 terms across all categories; 39% of terms).

Phase 3: Reconciling disagreement

6. MY, DG, LG independently articulated their personal descriptions or definitions used for each of the categories. Individual definitions of categories were submitted to MY and areas of similarities and differences across the category descriptions were identified. The articulation of category definitions was an extension of a previous team exercise where team members discussed individual definitions of clinical reasoning (Young et al 2018b).

7. A complete list of category descriptions, including identified areas of similarity and difference, were circulated to DG, MY and LG for review and discussion. Following discussion, it was decided that definitions for the six different categories were sufficiently similar for coding; however, this did not explain the lack of consensus for the remaining 67 terms. The analysis team decided that the lack of consensus in categorization was meaningful – that team members were likely interpreting the different terms and their inferred meaning differently rather than understanding the overarching categories differently. As a result, we engaged in one final round of open coding. If disagreements regarding inferred meaning of the terms persisted, the

team concluded that those disagreements were informative, rather than problematic. Rather than strive for team (all three agree) or majority-rules (two out of three agree) consensus, coding was considered to be complete following this final round.

8. A full data set describing: a) the terms for which there was consensus (n=43), b) the terms where two of the three members agreed (n=47), and c) the terms in which all three team members disagreed (n=20) was circulated to MY, DG and LG. The terms in which all three team members agreed were reviewed, discussed, and all agreed consensus was well founded. This round of coding focused on the terms for which two team members agreed (n=47). This step occurred via: a) MY, DG, LG independently reviewed the category descriptions generated by the analysis team. b) For the terms in which a given team member was in the 'minority' (i.e. two team members had assigned it to the same category, but different than the 'minority' member), the 'minority member' indicated whether they agreed with the majority code assignment or not (for n=31 terms, the team member who had assigned the 'minority code' agreed with majority coding; for n=16 terms team members stayed with their original coding). Following this coding, consensus was reached on a total of 74 terms.

During the analysis process, we decided that consensus coding of the entire list of 110 terms should not, and would not, be our goal. Therefore, data presented here are organized by: 1) those terms for which consensus was possible, 2) those terms for which majority consensus was possible (two of three coders agreed), and 3) terms for which no agreement across coders was documented.

Results

Thirteen coders identified 110 terms, across 625 papers, as being used to refer to clinical reasoning (Young et al [under review]). The 110 terms were identified a total of 693 times across the 625 papers.

Categories of terms and code descriptions:

The six categories of terms are described below and consensus terms associated with each category are presented in Table 1. The categories of terms, developed and refined through our coding process, were:

1. *Purpose/goal of reasoning* terms referred to the purpose, task, or goal for engaging in reasoning (e.g. patient management).

2. *Outcome of reasoning:* terms reflect a focus on the product or output of clinical reasoning or on a defined end-point of reasoning, with little emphasis on process or approach to reasoning (e.g. diagnosis).

3. *Reasoning performance:* terms reflect the mastery of performance or process(es) of reasoning (e.g. clinical competence, expert reasoning), or the quality of outcome of reasoning (e.g. judgment error).

4. Reasoning process: terms focus on the process components of clinical reasoning, encompassing processes, strategies, components, steps, or rationale used to solve a problem. Terms reflecting reasoning process tended to be tightly linked to a variety of cognitive or psychology-based frameworks (e.g. pattern recognition).

5. *Reasoning Skills:* terms imply clinical reasoning (or component processes) is stable and generalizable; a set of broad tasks or skills that are teachable, trainable, and/or by extension learnable (e.g. critical thinking).

6. *Context of reasoning:* This category reflects the influence of environmental, social, patient, or provider characteristics on clinical reasoning (e.g. situational awareness).

Categorization of terms used to refer to clinical reasoning:

The term, the categorization, and the frequency of mention (presented as percentage of all terminology mentions) are presented separately for:

1) *Terms for which complete consensus was possible* regarding the most appropriate categorization (n=74); see Table 1.

2) *Terms for which majority agreement* (defined as two of three coders) was possible, but no consensus was possible (n=16); see Table 2, and

3) *Terminology for which no agreement* or consensus was possible (n=20); see Table3.

4) Agreement regarding the *five most commonly used* terms to refer to clinical reasoning, and their categorization (where applicable), in Table 4.

[Insert Table 1]

[Insert Table 2]

[Insert Table 3]

[Insert Table 4]

Discussion:

This paper reports on the categorization of 110 terms used to refer to clinical reasoning, the degree of consensus possible regarding the inferred meaning of these terms across multiple rounds of coding, and the potential implications of the remaining disagreement for the teaching and assessment of clinical reasoning. Through an in-depth, iterative, and multi-stage coding process, we identified six main categories of terms: reasoning as a skill, reasoning as a process, reasoning as goal-oriented, reasoning as a means to an outcome, reasoning as contextually-influenced, or

reasoning performance. Though consensus was reached for the majority of terms (74 out of 110), we were unable to resolve differences in categorization for nearly one third of terms used to refer to clinical reasoning. For the five most commonly used terms in our database, consensus categorization was only possible for two of them (critical thinking and diagnostic reasoning), and no consensus was possible for the other three (decision making, problem solving, and clinical judgment), meaning that even some of the terms most commonly used to refer to clinical reasoning were understood differently by a team well versed in clinical reasoning, representing clinical educators and scholars. We decided that creating consensus through a negotiated common understanding of these terms was not as beneficial as documenting our disagreement. While speculative, it may be that some of these most common terms are used as broad substitutes for the equally broad construct of clinical reasoning, and, as such, tend to be interpreted differently depending on the focus, training, stance, teaching or assessment target, or even conceptualization of what constitutes clinical reasoning (Young et al 2018b). It may be that in addition to the term 'clinical reasoning' being a very broad concept that is differentially used (Young 2018) and differentially understood (Young et al 2018b), these common substitute terms may function as terminological Rorschach test – with individuals interpreting these terms through the lens of their particular understanding of clinical reasoning.

The lack of consensus across our team has important implications. It suggests that there is ambiguity in the inferred meaning of terms used to refer to clinical reasoning. For example, while we may assume that there is more specificity in 'clinical judgment' than 'clinical reasoning'; we did not agree on its categorization, and therefore its inferred meaning. In the context of teaching and assessment practices, this could mean that there may be inconsistent understanding of the terms being used to refer to clinical reasoning. Individuals – whether clinical educators, program

directors, or learners – may use the same terms with different intended meanings with potentially concerning pedagogical and assessment implications. For example, if the topic of a teaching session is on 'diagnostic judgment' – that concept could be understood differently; either as a reasoning purpose (the 'why' behind clinical reasoning), a reasoning skill (the 'habits' of good clinical reasoning), or as a reasoning performance (something to be mastered). Further, these different understandings of diagnostic judgement would likely result in very different assessment practices – one might focus on the ability to justify a decision with a formal rationale (assessing the 'why'), the other might examine how consistently one uses multiple sources of evidence to inform judgment (assessing the 'habits'), and another might examine whether there is documented evidence of improved accuracy of diagnosis across time (assessing mastery). Terms may also be differentially understood between teachers and learners – poteneially leading to confusion regarding what behaviours learners should be demonstrating to appear as competent reasoners (similar to procedural variation work done by Apramian et al (2015)).

The potential of multiple different understandings of a key term are not limited to clinical reasoning; for example, similar issues around different understandings of a construct have been identified in professionalism (Wynia et al 2014). In the case of clinical reasoning, confusion regarding the intended meanings of terms could be particularly problematic within competency frameworks if we assume a common understanding, or intended interpretation (*Accreditation Council for Graduate Medical Education Core Competencies* 2016, *the Royal College of Physicians and Surgeons of Canada's CanMEDS framework* 2015, *the General Medical Council's Good Medical Practice* 2014, *the Canadian Association of Occupational Therapists Profile of Practice* 2012, *the Canadian Physiotherapy Association Competency Profile* 2002), particularly

with a current focus on interprofessional collaborative practice and education (Interprofessional

Education Collaborative, 2016).

We want to reiterate that we intentionally did not seek to fully engage in consensus

methods. Though it may have been possible to use a Delphi, or adapted Delphi methodology, or to

impose a 'majority rules' categorization or coding rule, engaging in a process intended to support

consensus would have glossed over important disagreements brought to light by our analysis

process. Rather than trying to resolve disagreement, we decided to report it to facilitate a

discussion regarding differences in our understandings of the many terms for clinical reasoning.

One limitation may be that the larger review from which these terms were drawn focused on the

assessment of clinical reasoning, so our work may be missing important teaching-related

terminology for clinical reasoning. We have no evidence to support that language would be

different for publications focused on the teaching of clinical reasoning; however, 44% (n=277) of

publications in our corpus included an assessment of clinical reasoning (Young et al [under

review]).

Conclusion:

The data presented here suggest a diversity in understandings of a variety of terms used to refer to

clinical reasoning. Team members understood some terms similarly (e.g. cognitive bias,

diagnostic accuracy, adaptive expertise), but for others, team members inferred very different

meanings for the same term (e.g. judgment, decision making, clinical management), including

some of the most commonly used terms to refer to clinical reasoning. We believe that these

differences in understanding of the implied meaning behind a term are problematic for the

teaching and assessment of clinical reasoning; as teaching and assessment necessitate a similar

understanding, or at least operationalization, of a given construct. This work also suggests that

clinical reasoning is a multidimensional concept, with a wealth of association language, and a

multiplicity of both implied and inferred meanings, rather than a singular, unified, commonly

understood construct. We recommend caution when selecting terminology for the teaching and

assessment of clinical reasoning and suggest being more explicit when defining our intended

meaning of a given term, in order to support more clear communication of the expectations,

relevant targets, and desired outcomes of an educational or assessment encounter.

Practice Points:
1. Clinical reasoning is included within many of the current educational and assessment frameworks in health professions education.
2. There are a large number of terms used to refer to clinical reasoning, reflecting that clinical reasoning is likely a multifaceted construct.
3. Individuals understand the meaning of some of this terminology differently, suggesting that there may be room for confusion between educators, administrators, learners, and researchers.
4. Different understandings of clinical reasoning and associated terminology may make it difficult to clearly communicate educational and assessment objectives for a given encounter.
5. Being explicit regarding the intended meaning behind the language used should facilitate a common understanding within educational contexts across the health professions in order to better support teaching and assessment practices.
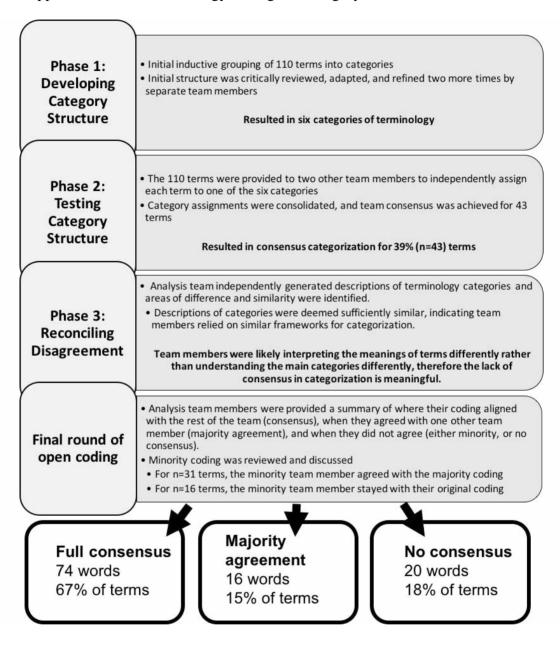
References:

Accreditation Council for Graduate Medical Education (ACGME). Common Program Requirements. 2016. Chicago. [June 10, 2017]. http://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/CPRs_07012016.pdf

Arpamian T, Cristancho S, Watling C, Ott M, Lingard L. Thresholds of principle and preference: exploring procedural variation in postgraduate surgical education. Acad Med. 90(11 Suppl):s70-s76.

Blumenthal-Barby JS, Krieger HA. 2015. Cognitive Biases and Heuristics in Medical Decision Making: A Critical Review Using a Systematic Search Strategy. Med Decis Mak. 35(4):539-557.

Canadian Association of Occupational Therapists. Profile of Practice of Occupational Therapists in Canada 2012. 2012. [June 10, 2017]. https://www.caot.ca/document/3653/2012otprofile.pdf

Canadian Alliance of Physiotherapy Regulators & Canadian Physiotherapy Association. Competency Profile: Essential Competencies for Physiotherapist Support Workers in Canada. 2002. Toronto: Authors.

Cutrer WB, Sullivan WM, Fleming AE. 2013. Educational strategies for improving clinical reasoning. Curr Probl Pediatr Adolesc Health Care. 43(9):248-257.

Eva KW. 2005. What every teacher needs to know about clinical reasoning. Med Educ. 39(1):98-106.

Frank JR, Snell, L, Sherbino J, editors. CanMEDS 2015 Physician Competency Framework. Ottawa:Royal College of Physicians and Surgeons of Canada. 2017. [June 10, 2017]. http://www.royalcollege.ca/rcsite/canmeds/canmeds-framework-e.pdf

General Medical Council. Good Medical Practice. 2014. Manchester: General Medical Council. [June 10, 2017]. https://www.gmc-uk.org/-/media/documents/Good_medical_practice___English_1215.pdf_51527435.pdf

Higgs J, Jones MA, Loftus S, Christensen N. (Eds.). 2008. Clinical reasoning in the health professions (3rd ed.).; Amsterdam: Elsevier Butterworth Heinemann.

Holmboe ES, Durning SJ. 2014. Assessing clinical reasoning: moving from in vitro to in vivo. Diagnosis. 1(1):111-117.

Ilgen JS, Humbert AJ, Kuhn G, Hansen ML, Norman GR, Eva KW, Charlin B, Sherbino J. 2012. Assessing diagnostic reasoning: a consensus statement summarizing theory, practice, and future needs. Kowalenko T, ed. Acad Emerg Med. 19(12):1454-1461.

Interprofessional Education Collaborative. Core competencies for interprofessional collaborative practice: 2016 update. Washington, DC: Interprofessional Education Collaborative.

Norman GR. 2005. Research in clinical reasoning: past history and current trends. Med Educ. 39(4):418-427.

Norman GR, Sherbino JD, Dore KL, Wood TJ, Young ME, Gaissmaier W, Kreuger S, Monteiro SD. 2014. The etiology of diagnostic errors: a controlled trial of system 1 versus system 2 processing. Acad Med. 89(2):277-284

Young ME, Dory V, Lubarsky S, Thomas AT. How different theories of clinical reasoning influence teaching and assessment. 2018a. Acad Med. 93(9):1415

Young ME. 2018. Crystallizations of constructs: lessons learned from a literature review. Perspect Med Educ. 7(Suppl 1):21-23.

Young ME, Thomas A, Lubarsky S, Ballard T, Gordon D, Gruppen L, Holmboe E, Ratcliffe T, Rencic J, Schuwirth L, Durning S. 2018b. Drawing boundaries: the difficulty of defining clinical reasoning. Acad Med. 93(7):990-995.

Young ME, Thomas A, Lubarsky S, Gordon D, Gruppen LD, Rencic J, Ballard T, Holmboe E, Da Silva A, Ratcliff T, Schuwirth L, Dory V, Durning SJ. Mapping clinical reasoning literature across the health professions. BMC Med Ed (Under review).

Wynia MK, Papadakis M, Sullivan WM, Hafferty FW. May 2014. More Than a List of Values and Desired Behaviors: A Foundational Understanding of Medical Professionalism. Acad Med. 89(5):712–714

Figure 1: Approach used for terminology coding and category refinement

Phase 1: Developing Category Structure
- Initial inductive grouping of 110 terms into categories
- Initial structure was critically reviewed, adapted, and refined two more times by separate team members

Resulted in six categories of terminology

Phase 2: Testing Category Structure
- The 110 terms were provided to two other team members to independently assign each term to one of the six categories
- Category assignments were consolidated, and team consensus was achieved for 43 terms

Resulted in consensus categorization for 39% (n=43) terms

Phase 3: Reconciling Disagreement
- Analysis team independently generated descriptions of terminology categories and areas of difference and similarity were identified.
- Descriptions of categories were deemed sufficiently similar, indicating team members relied on similar frameworks for categorization.

Team members were likely interpreting the meanings of terms differently rather than understanding the main categories differently, therefore the lack of consensus in categorization is meaningful.

Final round of open coding
- Analysis team members were provided a summary of where their coding aligned with the rest of the team (consensus), when they agreed with one other team member (majority agreement), and when they did not agree (either minority, or no consensus).
- Minority coding was reviewed and discussed
  - For n=31 terms, the minority team member agreed with the majority coding
  - For n=16 terms, the minority team member stayed with their original coding

Full consensus
74 words
67% of terms

Majority agreement
16 words
15% of terms

No consensus
20 words
18% of terms

Table 1: Terms for which consensus was possible (n=74 terms):
Note: Proportion of mentions is the number of times a given term was reported from our corpus given total mentions (number of mentions divided by 693)

Major code	Term	Proportion of mentions
Purpose/goal of reasoning	Diagnostic reasoning	6.93%
	Clinical problem-solving	5.63%
	Diagnostic thinking	0.87%
	Diagnostic decision making	0.58%
	Treatment decision making	0.29%
	Case management	0.14%
	Clinical identification	0.14%
	Diagnostic justification	0.14%
	Patient Management	0.14%
	Surgical decision making	0.14%
	Therapeutic reasoning	0.14%
Outcome of reasoning	Diagnosis	1.15%
	Diagnostic Accuracy	1.15%
	Differential diagnosis	0.43%
	Premature closure	0.29%
	Accuracy	0.14%
	Choice of treatment	0.14%
	Classification (classification of cardiogram tracings)	0.14%
	Clinical management decisions	0.14%
	Cognitive bias	0.14%
	Diagnostic and management quality	0.14%
	Diagnostic success	0.14%
	Error prevention	0.14%
	Judgement errors	0.14%
	Medical error	0.14%
	Reasoning errors	0.14%
	Management plan	0.14%
Reasoning Performance	Competency	1.01%
	Diagnostic performance	0.87%
	Clinical competence	0.72%
	Clinical performance	0.29%
	Diagnostic expertise	0.29%
	Adaptive expertise	0.14%
	Cognitive expertise	0.14%
	Diagnostic acumen	0.14%
	Expert reasoning	0.14%
	Expertise	0.14%
	Medical expertise	0.14%

Reasoning processes	Hypothetico-deductive reasoning	0.72%
	Bayesian probabilistic thinking	0.58%
	Intuition	0.58%
	Pattern recognition	0.58%
	Inductive and deductive reasoning	0.43%
	Analytic reasoning	0.29%
	Analytical thinking	0.14%
	Backward forward reasoning	0.14%
	Backward reasoning	0.14%
	Cognitive processes	0.14%
	Critical appraisal	0.14%
	Enabling conditions	0.14%
	Heuristics	0.14%
	Higher order thinking	0.14%
	Intuitive reasoning	0.14%
	Medical information processing	0.14%
	Pattern matching	0.14%
	'Street diagnosis' or in the blink of the eye	0.14%
	Metacognition	0.14%
	Reasoning strategies	0.14%
	Reflective thinking skills	0.14%
	Self monitoring	0.14%
Reasoning skills	Critical thinking	15.73%
	Clinical skills	0.87%
	Cognitive skill	0.14%
	Critical analysis	0.43%
	Critical reasoning	0.14%
	Reasoning	0.14%
	Reasoning skills	0.14%
Context of reasoning	Situational judgement	0.43%
	Situation awareness	0.29%
	Dialectical reasoning	0.14%
	Informed decision making	0.14%
	Participatory decision making	0.14%
	Shared decision making	0.14%
	Shared understanding	0.14%

Table 2: Terminology for which majority categorization was possible (n=16):
Note: proportion of mentions is the number of times the term was identified in our corpus divided by the total mentions (n=693).

Term	Majority (2 of 3 coders) categorization	Alternate categorization	Proportion of mentions
Problem solving	Purpose/goal of reasoning	Reasoning skills	13.28%
Nursing process	Context of reasoning	Reasoning skills	0.29%
Expert judgment	Reasoning Performance	Purpose/goal of reasoning	0.29%
Conditional reasoning	Purpose/goal of reasoning	Reasoning processes	0.14%
Narrative reasoning	Purpose/goal of reasoning	Reasoning processes	0.14%
Practical reasoning	Purpose/goal of reasoning	Reasoning processes	0.14%
Pragmatic reasoning	Purpose/goal of reasoning	Reasoning processes	0.14%
Triage decisions	Purpose/goal of reasoning	Outcome of reasoning	0.14%
Professional judgement	Reasoning performance	Reasoning skills	0.14%
Diagnostic processes	Reasoning processes	Purpose/goal of reasoning	0.14%
Discourse organization	Reasoning processes	Outcome of reasoning	0.14%
Hypothesis testing	Reasoning processes	Purpose/goal of reasoning	0.14%
Lines of reasoning	Reasoning processes	Reasoning performance	0.14%
Logic	Reasoning processes	Purpose/goal of reasoning	0.14%
Clinical thinking	Reasoning skills	Purpose/goal of reasoning	0.14%
Nursing reasoning	Reasoning skills	Context of reasoning	0.14%

Table 3: Terminology for which all coders disagreed (n=20): *Note: proportion of mentions is the number of times the term was identified in our corpus divided by the total mentions (n=693). Those with the same proportion of mentions are presented alphabetically.*

Term	Categorization by Rater 1	Categorization by Rater 2	Categorization by Rater 3	Proportion of mentions
Decision making	Outcome of reasoning	Purpose/goal of reasoning	Reasoning skills	21.79%
Clinical judgement	Reasoning skills	Reasoning performance	Purpose/goal of reasoning	8.66%
Judgement	Reasoning skills	Reasoning performance	Purpose/goal of reasoning	2.16%
Management	Outcome of reasoning	Reasoning skills	Purpose/goal of reasoning	0.43%
Surgical judgement	Reasoning skills	Reasoning performance	Purpose/goal of reasoning	0.43%
Clinical assessment	Outcome of reasoning	Reasoning skills	Unable to categorize	0.29%
Knowing	Unable to categorize	Context of reasoning	Reasoning processes	0.29%
Application of knowledge	Reasoning skills	Reasoning performance	Reasoning processes	0.14%
Cause effect reasoning	Reasoning skills	Reasoning processes	Purpose/goal of reasoning	0.14%
Clinical management	Outcome of reasoning	Reasoning skills	Purpose/goal of reasoning	0.14%
Clinical perception	Purpose/goal of reasoning	Reasoning processes	Unable to categorize	0.14%
Conceptual thinking	Reasoning skills	Reasoning processes	Purpose/goal of reasoning	0.14%

Diagnostic judgement	Reasoning skills	Reasoning performance	Purpose/goal of reasoning	0.14%
Error recovery	Reasoning processes	Outcome of reasoning	Purpose/goal of reasoning	0.14%
Evidence-based clinical decision-making	Purpose/goal of reasoning	Reasoning processes	Reasoning skills	0.14%
Interactive reasoning	Purpose/goal of reasoning	Context of reasoning	Unable to categorize	0.14%
Knowledge application	Reasoning skills	Reasoning performance	Purpose/goal of reasoning	0.14%
Lateral thinking	Reasoning skills	Reasoning processes	Unable to categorize	0.14%
Psychodiagnosis	Outcome of reasoning	Reasoning skills	Purpose/goal of reasoning	0.14%
Synthesis	Outcome of reasoning	Purpose/goal of reasoning	Reasoning processes	0.14%

Table 4: Five most commonly mentioned terms to refer to clinical reasoning:

Term	Proportion of mentions	Code	Consensus coding possible?
Decision making	21.8%	-	no
Critical thinking	15.73%	Reasoning skills	yes
Problem solving	13.2%	-	no
Clinical judgement	8.6%	-	no
Diagnostic reasoning	6.93%	Purpose/goal of reasoning	yes