Article:
Lee, M.J., Lamidi, S., Williams, K.M. et al. (2024) Commentary: core descriptor sets using consensus methods support 'table one' consistency. Journal of Clinical Epidemiology, 174. 111470. ISSN: 0895-4356

https://doi.org/10.1016/j.jclinepi.2024.111470

Journal of Clinical Epidemiology 174 (2024) 111470

Journal of Clinical Epidemiology

Commentary: core descriptor sets using consensus methods support 'table one' consistency

Matthew J. Lee[a,*], Segun Lamidi[b], Kate M. Williams[c], Sue Blackwell[a], Adil Rashid[d], Peter O. Coe[e], Nicola S. Fearnhead[f], Natalie S. Blencowe[g,h], Daniel Hind[i]

[a]*Institute for Applied Health Research, University of Birmingham, Birmingham, UK*
[b]*Department of Cardiovascular Sciences, University of Leicester, Leicester, UK*
[c]*Barnsley Hospital, Barnsley NHS Foundation Trust, Barnsley, UK*
[d]*Clinical Effectiveness Unit, Royal College of Surgeons of England, London, UK*
[e]*Department of Upper Gastrointestinal Surgery, Leeds Teaching Hospitals, Leeds, UK*
[f]*Department of Surgery, Addenbrookes Hospital, Cambridge, UK*
[g]*Centre for Surgical Research, Population Health Sciences, Bristol Medical School, Bristol, UK*
[h]*Department of Emergency General Surgery, Leeds Teaching Hospitals, Leeds, UK*
[i]*Clinical Trials Research Unit, Section of Health and Related Research, University of Sheffield, Sheffield, UK*

Accepted 15 July 2024; Published online 20 July 2024

Abstract

Background: Inconsistent reporting of patient characteristics in clinical research hampers reproducibility and limits analysis opportunities. This paper proposes condition-specific 'Core Descriptor Sets' comprising key factors like demographics, disease severity, comorbidities, and prognosis to standardize Table 1 reporting.

Methods: Development entails stakeholder involvement, systematic identification of descriptors, value rating, and consensus-building using multiple Delphi rounds. Final agreement comes at an expert meeting.

Conclusion: Benefits include easier cross-study comparison, for example, through individual patient meta-analysis, facilitated by comparison of consistently reported individual data rather than group-level analysis. This may also support routine data analyses, subgroup and risk identification, and reduced research waste. Core Descriptor Sets describe cohorts thoroughly while minimizing research burden. They are intended to enable improved clinical characterization, personalization, reproducibility, data sharing, and knowledge building.

Keywords: Consensus; Prognosis; Research design; Risk factors; Data analysis; Documentation; Epidemiologic research design; Guidelines as topic; Publishing; Reproducibility of results

1. Introduction

Consistency of reporting has been a focus for clinical research in recent years, with efforts made to improve the reporting of interventions and outcomes [1,2]. There remains significant variation in how we define the population and disease being studied. There is currently much variation in how this is reported across a range of conditions (Table one) [3–6]. This poses downstream problems for generalizability and evidence synthesis [6]. How can we ensure that we describe the situation at baseline in a manner that adequately describes the patient demographics, comorbidities, and other prognostic features, as well as capturing disease severity and impact?

In clinical studies, patient and disease characteristics are summarized in 'Table One', which serves rhetorical functions related to external and internal validity [7]: selection bias may affect treatment estimates; disease severity and comorbidities should be balanced at baseline; and the study population should resemble the target population [8–10]. ''Table One'' offers insight into implicit theories of disease or equity, where the characteristics presented are assumed to modify treatment effects and/or indicate inclusivity. Stratified and personalized medicine has developed through understanding the disease course, prognosis factors,

Funding: No specific funding supported this work.

* Corresponding author. Institute for Applied Health Research, University of Birmingham, Birmingham B15 2TT, UK.
E-mail address: m.j.lee.1@bham.ac.uk (M.J. Lee).

https://doi.org/10.1016/j.jclinepi.2024.111470

What is new?

Key findings

- Consensus derived population of table one may help to standardise reporting of patient and disease characteristics.

What this adds to what was known?

- This piece demonstrates the philosophical and scientific justification and method to agree common baseline descriptors in clinical studies.
- This provides a framework to agree on core descriptors of patients and disease at baseline in studies, drawing on community opinions.

What is the implication and what should change now?

- Core descriptor sets should be developed for common conditions, and implemented into clinical research and practice.
- This will allow better understanding of population and disease factors that influence outcomes.

predictive modelling, and tailored treatment [11]. However, some areas have been slow to develop stratified or personalized care [12].

The 'reference class problem,' which relates to calculating individual risk from population data, posits that ''every single thing or event has an indefinite number of properties or attributes observable in it, and might therefore be considered as belonging to an indefinite number of different classes of things'' [13]. In the same way, philosophers recognize the concepts of our inquiry—including population descriptors—as systemically and dynamically open, with fuzzy or flexible—rather than crisp or rigid—boundaries [14]. They are 'ontologically multiple'—in that the entities we encounter are not singular or fixed but rather involve multiple, diverse, and sometimes conflicting ways of understanding and representing them [15]. Appeals to context are infinitely extendible [16]. As a group of researchers, we should attempt to achieve (at least) temporary consensus on the concepts of our inquiry. This consensus will likely trade-off long-term theory development against near-term practical concerns about utility and burden, and the impulse to unify against the impulse to distinguish concepts [17].

We propose a standardized, problem-specific, approach to enable easier cross-study comparisons of the contents of ''Table 1'': the 'Core Descriptor Set' (CDS), a list of key patient/disease characteristics, including factors important for classification, disease severity, or prognosis. Our approach is informed by the tenets of philosophical pragmatism [18] in that descriptor sets are:

- Problem-orientated: offering practical solutions to inconsistent baseline patient descriptor reporting;
- Antifoundational: presenting our knowledge of populations as contingent and evolving;
- Community-based: developed democratically in a community of inquiry comprising patients and clinicians;
- Situationally and context aware: tailored to specific conditions, concerned with the tension between the impulse for ''unification'' and for ''distinction'' (or ''lumping'' and ''splitting''), as well as between;
- Consequentialist: focused on the practical difference, relevance, and utility of choosing certain descriptors over others, in terms of treatment decisions and outcomes; and
- Fallibilist: in that descriptor sets remain open to revision pending new evidence; they represent provisional, warranted assertions, not final truths.

A CDS is intended to be a list of patient and disease characteristics that can be reported in all research studies

Table 1. Examples of baseline descriptor challenges

Author (y)	Disease of interest	Commentary on findings
Khan (2024)	Crohn's anal fistula	Median number of descriptors is 15. Some commonality seen within the procedural and medical groups, but these differ from each other. Some significant differences noted in reporting between groups, which might relate to key prognostic factors,e.g anorectal stenosis, proctitis.
Rashid (2024)	Adhesive Small Bowel Obstruction	Identified 156 descriptors across 73 studies. Median number of descriptors is 12, ranging from 1 to 34. Some consistency noted in reporting of age, sex, body mass index, cause of bowel obstruction, prior surgery, and American Society of Anesthesiologists classification.
Lee (2020)	Perforated peptic ulcer	Identified 76 descriptors across 23 studies. Median number of descriptors was 8, ranging from 3 to 13. Demographics generally reported, but poor reporting of disease risk factors and severity, and baseline physiology.
Wertli (2013)	Lower back pain	Review showed good reporting of general characteristics, but poor reporting of psychosocial factors, clinical examination findings, and baseline health status. This was felt to impede generalizability.

Table 2. The DECODE (DEveloping COre DEscriptors) checklist

Domain	Items to report	Rationale
Specification of condition and population to be described	• Condition(s) or subpopulations of interest, e.g. diabetes mellitus with peripheral neuropathy. • Acuity of condition/setting care provided in, e.g. emergency hospital presentation. • Specify geography if relevant (i.e. if specific to a country or region) • Specify other general characteristics as appropriate, e.g. paediatric patients	Ensures that the final CDS can be applied to the clinically important population of interest
Management group	• Number of participants and associated clinical or methodological expertise • Named lay representative	Demonstrate relevant range of expertise in the topic and method
Long list generation	• Method used to generate long list • Number of items identified in the long list	Provides insight into prior coverage of descriptors in the field
Patient involvement	• Documentation of feedback on long list, including items added • Feedback on final CDS including acceptability and perceived burden	Ensures relevant symptom descriptors are included, and that the final CDS is acceptable to future research participants.
Delphi participant recruitment	• How approaches were made to participants • Report of representativeness of participants including background specialism and geography.	Reflects how well participants reflect the general clinical population
Delphi voting process	• Prespecified thresholds for inclusion/exclusion • Number of items assessed at each stage	Demonstrates robustness of reporting of Delphi voting
Final consensus process	• Number of participants, clinical specialism(s), and geography. • Number of items rated • Number of items included in final CDS	Ensures participants are drawn from across the stakeholder group, and supports robust reporting of the final CDS.

CDS, Core Descriptor Set.

for a defined clinical condition. The descriptor set should include those key items which clinicians believe are important to classification of patient, disease, or prognosis. There is no fixed number of items to be included, but studies typically report 3–30 descriptors in Table One [4]. Researchers should undertake a pragmatic assessment of the trade-off between detail on multiple characteristics, and overloading readers [19], alongside the necessity for larger sample sizes in prognostic modelling [20].

2. Methods

A CDS development process should include, and report on, scope definition, collation of existing knowledge, and a consensus process (Table 2).

2.1. Research team, steering group, and patient involvement

The team should include patient representatives, health professionals with experience in managing the condition, and researchers with experience in systematic review and consensus methods. They should engage co-researchers from other health systems to maximize generalizability. Meaningful Patient and Public Involvement (PPI) should be reported as it is integral to improving the quality and relevance of research, through patient perspectives on the utility of research design, conduct, analysis, and dissemination. An external steering group should be convened with relevant knowledge of the clinical and methodological challenges, as well as lay representation. This steering committee will support the research team in ensuring robust methodology with engagement of stakeholders.

In previously completed studies, lay members felt unable to weigh all descriptors for the entire cohort given varied patient journeys [21]. Hence, PPI focused on integrating patients in CDS design, finalization, and presentation rather than the Delphi consensus itself, contrasting with core outcome set development. This position requires clarity upfront while affirming PPI's integral role.

The CDS development process should seek broad input to maximize generalizability. However, the scope of an individual CDS may need to be narrowly defined to ensure relevance and feasibility for a specific research area. Researchers should, where appropriate, gather perspectives across specialties, professions, health systems, and countries. The stakeholder network should reflect diverse views on the importance of descriptors for patients with the condition. This approach helps build consensus on the most essential items for concise yet comprehensive reporting.

2.2. Scoping

Careful scope definition should occur prior to study commencement. A CDS should be concise and relevant, minimizing researcher, and research participant burden. Broad population descriptors, encompassing numerous subgroups, may lack relevance. Limiting scope, to specific clinical procedures, conditions, or acuity, will increase clarity. Patients should help define the condition.

2.3. Identification of candidate items

An initial long list of descriptors can be extracted from the baselines tables of trials, and prognostic factors, identified by systematic reviews [21], as well as through bespoke descriptor reviews [3,4]. Communities may decide that items represent only 'what' should be measured, for example, 'mortality risk estimate,' 'pain,' and 'quality of life.' They should avoid being prescriptive about 'how,' for example, specifying particular psychometrically validated instruments or tools. They may however specify approaches such as computed tomography or ultrasound scanning. They should record decisions to decompose items from composite instruments, for instance 'quality of life,' into component domains or items, such as 'anxiety' or 'activities of daily living.' Patients should help review the initial long list, providing additional descriptors related to symptoms or social determinants that influence decisions but are under-represented in the literature.

The research group reviews the long list, consulting other settings if limited external input. Review involves all stakeholders, via focus groups or similar approaches, to separate inadvertent composite measures and include missing descriptors particularly relevant in defining the population. Lay members provide critical input on descriptors especially important to patients.

2.4. Iterative delphi process

The long list likely contains more items than can be reported in a paper. Consensus building with stakeholders using an online Delphi exercise over 3 voting rounds helps prioritize key descriptors. Participants consider an item's importance in describing patients and influencing treatment decisions or outcomes. Round 1 presents items randomly with 1-9 Likert scale ratings plus suggestions for missing items. Preset thresholds determine if an item carries forward.

2.5. Consensus meeting

The exercise yields a core set of descriptors for the condition. The group avoids recommending measurement tools; selected measures should have reliability and validity, requiring separate evaluation. This process brings key stakeholders together to determine essential items for concise yet comprehensive reporting.

A consensus meeting with diverse stakeholders reviews the final item list. Researchers may present items which are at the borderline of acceptance (i.e within 1%−2% of threshold for inclusion in round 3) if late Delphi rounds had small numbers of raters and:

- Review included and borderline items;
- Vote on borderline items; and
- Highlight exclusion discrepancies (eg, by clinical specialty) [3].

Teams must seek patient feedback on acceptability, appropriateness, burden, and feasibility, of data collection, including at critical points in the patient pathway, such as diagnosis and when management strategies change, and for underserved groups, including those with limited language proficiency. Domains describing the internal structure of the data (see above) are presented for validation. Where shortlisted items are considered excessive, participants may prioritize domain items.

3. Discussion

The value of ideas derives from their practical consequences [18]. We have proposed an approach to 'Table One' descriptors based on consensus about their importance for treatment decisions and for understanding outcomes. Comprehensive description and classification may improve recognition of subgroups for stratification or personalization of care [11]. In general, we could think of Table One data, not as referring to static entities, but as something on which we continuously reflect, and which we transform [22]. CDSs may aid the identification of borderline cases unaddressed by prevailing clinical models shaped by investigator attitudes [22], enabling newer classification approaches [23]. It is important to note that standardizing descriptors goes beyond Table One, and has implications for overall study design, particularly in determining which baseline characteristics to measure. The CDS aims to provide a starting point for improving consistency in these areas.

Our approach advocates collective experimentation [18]—clinicians and patients from diverse settings addressing problems encountered in experience through the exploration of their gestalt preferences. Exploring descriptor preferences may reveal, and enable us to challenge, habitual assumptions and entrenched ways of thinking about clinical entities, which are rarely made explicit. This method appears to be of interest to researchers and clinicians as 3 papers have been published using this method [21,24,25], with studies in venous disease, and solid organ transplantation in development.

As pragmatists we are also concerned with collective sense-making [18]. We hope that standardizing descriptors through a core set should enable more effective aggregation and comparison of data across studies in meta-analysis of

epidemiological studies [26]. Increasing specialization of medicine and biomedical research may make it difficult to connect clinical decision-making to basic science. If we recognize an understanding of basic physiological mechanisms as necessary for decision-making, we can benefit from placing epidemiological research in networks of theoretical models. Consequently, triangulation between different study types might yield more robust research and more significant advances [27,28].

Science faces challenges around reproducibility and efficiency. A community level adoption of a standardized approach to the construction of 'Table One' may lead to efficiencies of design and data collection, provide context for heterogeneity in meta-analysis, and support the understanding of disease and classification. This could improve our understanding of disease, and help us better stratify patients by characteristics, moving closer to stratified medicine.

CRediT authorship contribution statement

M.J. Lee: Writing − review & editing, Writing − original draft, Supervision, Project administration, Methodology. **S. Lamidi:** Writing − review & editing, Methodology. **K.M. Williams:** Writing − review & editing, Methodology. **S. Blackwell:** Writing − review & editing, Methodology. **A. Rashid:** Writing − review & editing, Methodology. **P.O. Coe:** Writing − review & editing, Methodology. **N.S. Fearnhead:** Writing − review & editing, Methodology, Conceptualization. **N.S. Blencowe:** Writing − review & editing, Methodology, Conceptualization. **D. Hind:** Writing − review & editing, Writing − original draft, Supervision, Methodology, Conceptualization.

Data availability

No data was used for the research described in the article.

Declaration of competing interest

None to declare.

References

[1] Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: Template for intervention description and replication (TIDieR) checklist and guide. Gesundheitswesen 2016;78:e174.
[2] Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, et al. Core outcome set-STAndards for development: the COS-STAD recommendations. PLoS Med 2017;14:e1002447.
[3] Khan S-Z, Arline A, Williams KM, Lee MJ, Steinhagen E, Stein SL. The use of core descriptors from the ENiGMA code study in recent literature: a systematic review. Colorectal Dis 2024;26:428−38.
[4] Lee MJ, Coe PO, O'Donoghue R, Peirson M, Saha A. Variation in descriptors of patient characteristics in randomized clinical trials of peptic ulcer repair: a systematic review. Br J Surg 2020;107:1570−9.
[5] Rashid A, Kumar M, Lee MJ. A systematic review of participant descriptors reported in studies of adhesive small bowel obstruction. Colorectal Dis 2024;26:851−70.
[6] Wertli MM, Schöb M, Brunner F, Steurer J. Incomplete reporting of baseline characteristics in clinical trials: an analysis of randomized controlled trials and systematic reviews involving patients with chronic low back pain. PLoS One 2013;8:e58512.
[7] Hayes-Larson E, Kezios KL, Mooney SJ, Lovasi G. Who is in this study, anyway? Guidelines for a useful Table 1. J Clin Epidemiol 2019;114:125−32.
[8] Bijkerk CJ, Muris JWM, Knottnerus JA, Hoes AW, de Wit NJ. Randomized patients in IBS research had different disease characteristics compared to eligible and recruited patients. J Clin Epidemiol 2008;61:1176−81.
[9] Horwitz RI, Lobitz G, Mawn M, Conroy AH, Cullen MR, Sim I, et al. Rethinking table 1. J Clin Epidemiol 2022;142:242−5.
[10] Nguyen T-L, Xie L. Incomparability of treatment groups is often blindly ignored in randomised controlled trials - a post hoc analysis of baseline characteristic tables. J Clin Epidemiol 2021;130:161−8.
[11] Hemingway H, Croft P, Perel P, Hayden Ja, Abrams K, Timmis A, et al. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. BMJ 2013;346:e5595.
[12] Mukherjee R. An acute need: precision medicine for acute care surgery. BJS Open 2023;7:zrad003.
[13] Venn J. The Logic of Chance: An Essay on the Foundations and Province of the Theory of Probability, With Especial Reference to Its Logical Bearings and Its Application to Moral and Social Science, and to Statistics. 3rd ed. London: Macmillan; 1888.
[14] Kaplan A. Concepts. The Conduct of Inquiry: Methodology for Behavioral Science. San Francisco: Chandler Publishing Company; 1964:34−83.
[15] Mol A. The Body Multiple: Ontology in Medical Practice. London: Duke University Press; 2002.
[16] Scharfstein B-A. The Dilemma of Context. New York: NYU Press; 1989.
[17] James W. The sentiment of rationality. Mind 1879;4:317−46.
[18] Wills J, Lake R. The power of pragmatism: knowledge production and social inquiry. Manchester, UK: Manchester University Press; 2020.
[19] Pierson E, Goodman N. Uncertainty and denial: a resource-rational model of the value of information. PLoS One 2014;9:e113342.
[20] Riley RD, Ensor J, Snell KIE, Harrell FE Jr, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020;368:m441.
[21] ENiGMA CODE collaborators. Development of a core descriptor set for Crohn's anal fistula. Colorectal Dis 2023;25:695−706.
[22] Bowker GC, Star SL. Sorting things out: classification and its consequences. Boston, MA: MIT Press; 2000.
[23] Mori M, Krumholz HM, Allore HG. Using latent class analysis to identify hidden clinical phenotypes. JAMA 2020;324:700−1.
[24] Lamidi S, Williams KM, Hind D, Peckham-Cooper A, Miller AS, Smith AM, et al. Defining core patient descriptors for perforated peptic ulcer research:: international Delphi. Br J Surg 2022;109:znac096.
[25] PelvEx Collaborative. Empty pelvis syndrome: PelvEx Collaborative guideline proposal. Br J Surg 2023;110:1730−1.
[26] Baker WL, White CM, Cappelleri JC, Kluger J, Coleman CI, Health Outcomes, Policy, and Economics (HOPE) Collaborative Group. Understanding heterogeneity in meta-analysis: the role of meta-regression. Int J Clin Pract 2009;63:1426−34.
[27] Hampton T. Bench to bedside and back again may Be key to clinical breakthroughs. JAMA 2017;318:16−7.
[28] Uzzi B, Mukherjee S, Stringer M, Jones B. Atypical combinations and scientific impact. Science 2013;342:468−72.