

Critical Appraisal on the Selection Criteria for Evidence-based Healthcare

Paulo Cesar Morales Mayer

Universidade Ceuma <https://orcid.org/0000-0001-9484-2118>

Caroline Amélia Gonçalves

Universidade Ceuma

Franz Porzsolt (✉ mindset@clinical-economics.com)

Institute of Clinical Economics (ICE) e.V., D-89081 Ulm, Germany. <https://orcid.org/0000-0003-3554-2902>

Research article

Keywords: eligibility criteria, selection criteria, evidence-based, exclusion criteria, spectrum bias

Posted Date: January 15th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-145261/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Evidence-Based healthcare deals basically with published clinical trials to guide the decision making on what treatment to use for any specific conditions.

Aims: The present paper assessed the inclusion and exclusion criteria used in clinical trials of cervical cancer aiming at establishing a clear distinction between each criterion.

Methods: We performed a bibliographical search in pubmed with the terms cervical cancer and treatment or therapy filtered for clinical trials with human subjects for the last ten years. A total of 30 papers were used extracting and classifying the inclusion and exclusion category according to the characteristic they described.

Results: We found no clear parameter to establish which criteria could exclusively serve as inclusion or exclusion across the papers, about 56% of the categories identified were found either listed as inclusion or exclusion criteria or even as both in some cases.

Conclusions: The key issue of selection criteria is not in its form but in its function, the first point to consider is if the trial is experimental (focused on efficacy and proof of principle) or observational (pragmatic trials, focused on effectiveness and real world conditions). We suggest, inclusion criteria should be broad, focused on the investigated condition; exclusion criteria should apply only to the subset of this "included" population, and do not take part in observational studies. These conclusions do not serve only for researchers but should affect practitioners and policy makers to correctly compare the results of investigated treatment.

Background

Systematic reviews and metanalysis are the most important elements for evidence-based-medicine (1). The number of these studies have been increasing across disciplines since 2006, it is estimated that more than 250,000 systematic reviews have been published so far (2). Although the studies differ in quality and reliability (3) there are several initiatives, tools and protocols not only to perform but also evaluate the reviews (e.g. PRISMA, GRADE, PROSPERO, Cochrane, AMSTAR, JADAD scale, ROBINS tool to assess the risk of bias, and so on) and a continuous effort to refine them as well (4).

Despite these efforts, systematic reviews are still hostage of its raw material, the clinical trials themselves. For an adequate comparison of the results of multiple trials it is essential they are equivalent, not only in its methods (the usual concern in a review) but also in the population the treatment was tested. Sometimes called eligibility or selection criteria the parameters for the participant assignment to the study are described in the methods as inclusion and exclusion criteria. However even research published in ranked journals sometimes does not provide clear description of the criteria for the selection of the participants.

Comparing multiple studies with unclear selection criteria of participants may result in a problem known as “spectrum bias” (5) i.e., differences in sensitivity and specificity of the treatment due to a different range of participant characteristics. It is known since the early 80’s that the spectrum of similar patient groups may differ considerably (6). A literature search using the term “spectrum bias” identified about 500 publications. Usually, two types of criteria are applied to select the appropriate population of patients. These two types are the inclusion and exclusion criteria. Several papers discuss the adequacy or specific inclusion/exclusion criteria for particular conditions, demonstrating that there is some scientific concern on the subject (7,8).

A previous study reported a high variation in the numbers of inclusion and exclusion criteria in 75 clinical trials. The median number of reported selection criteria was 6 with an interquartile range (IQR) of 8 - 10. The reported number was in 15% of studies lower than the IQR and in 21% of studies higher. Inclusion (exclusion) criteria were not reported in 25% (31%) of studies (9).

However, there is no clear differentiation between these criteria or even these differential functions in many clinical trials, even though a precise description of them is important not only to identify the homogeneity/heterogeneity of the population studied, but also to assess the very internal and external validity of the study and to identify which studies are actually comparable. Despite the difference in objectives, variables, methods and conditions, across studies, is it possible to establish a general clear distinction between inclusion and exclusion criteria?

The objectives of the present study were to: (1) assess the use and differentiation of inclusion/exclusion criteria in clinical trials of a particular condition, and (2) propose a clear distinction between inclusion and exclusion criteria.

Methods

This study consisted in two phases, I – categorization and analysis of the selection criteria and II – verification of the classification in other studies.

For phase I, we performed a search at MEDLINE (via- pubmed) using the search terms in title and abstract: “Treatment” with the operators OR “Therapy”, AND “Cervical Cancer”. The search strategy corresponded to the following command: ((therapy[Title/Abstract]) OR treatment[Title/Abstract]) AND cervical cancer[Title/Abstract]. Also, the following filters were applied: Clinical Trials, Free Full Text, Last ten years.

After reading the title and the abstract, were included in the study only primary studies accessing treatments for cervical cancer (i.e. safety, efficacy and/or effectiveness). Were excluded from the analysis, reviews, pre-clinical, or *in vitro* studies, protocols, economical value studies, and studies focusing on secondary effects of the treatment (e.g. quality of life, specific physiological functions).

After downloading the included studies, the first and the second authors used an extraction spreadsheet, retrieving the following information from each paper independently: paper reference, inclusion criteria, exclusion criteria. After performing this procedure for five papers, we created categories to classify each of the criteria and plotted in a chart indicating in which study each category was present. We applied the categories to the next five papers assessing its adequacy to fit the inclusion and the exclusion criteria, and the need to create additional categories. This part of the study was concluded after three consecutive rounds (15 papers) of perfect fit of the categories without the need for additional categories.

The categories were classified among: inclusion, exclusion, or both inclusion and exclusion, according to the description on the paper.

For phase II we performed another search in MEDLINE to identify if the exclusively inclusion, and exclusion criteria were actually exclusive for this selection criteria in a general search in other clinical trials. We used terms related to each category created in phase I as search terms, restricting the search to clinical trials.

Results

A total of 18 categories were identified (Table 1). The last new category was added from the 14th paper, applying the set of criteria of three consecutive rounds of five papers with no additional category, the paper selection was interrupted with 30 papers (Figure 1). The maximum number of added categories by a single paper was six.

From Figure 1 it can be noticed that all of the 30 papers mentioned the inclusion criteria, seven papers mentioned only inclusion criteria, 13 papers mentioned categories exclusively in inclusion or exclusion criteria, and 10 papers mentioned at least one category (range 1-3 categories) both as inclusion and exclusion criteria.

From Table 2 it can be noticed that out of the 18 categories five were exclusively mentioned in inclusion criteria, three only as exclusion criteria and the remaining 10 categories were used both as inclusion or exclusion criteria. No category was present in all the papers (however the investigated condition was implicit in other parts of the paper when not directly mentioned). The most frequent categories for inclusion criteria were: "Investigated condition" and "Standardized performance status" present in 27 and 26 papers respectively. The next most frequent categories mentioned as inclusion criteria were "Disease staging or characteristic", "Physiologic functions" and "Age", including the papers which mentioned them in both criteria. For exclusion criteria the most frequent categories were, "Described comorbidities", "Treatment" and "Medical History", including the papers which mentioned them in both criteria as well. The categories used only for exclusion criteria were the least frequent, present in 1 to 3 papers only.

Five categories ("Described comorbidities"; "Disease staging or characteristic"; "Physiologic functions"; "Pregnancy/breastfeeding"; "Treatment") appeared in at least two papers (ranging from 2-6) both in inclusion and exclusion criteria.

Analyzing the papers in which the same category was used both as the inclusion and exclusion criteria we identified circumstances in which it was justifiable and others in which it could be avoidable. Justifiable were those situations in which some characteristics of the category would be needed and others would compromise the evaluation of the intervention, for example, if the treatment was for early stages of cervical cancer, exposure to some previous treatment (e.g. tumor removal surgery) would be necessary to evaluate the efficacy of some adjuvant therapy to prevent relapse (inclusion) while exposure to chemotherapy could confound the results of the intervention tested (exclusion). Avoidable were those situations in which there is a redundancy of the criteria, for example, negative beta HCG test (inclusion) and pregnancy (exclusion), or negative description of different characteristic, for example, absence of infectious disease (inclusion) and presence of autoimmune disease (exclusion).

As can be observed from Figure 2, in 15 occasions the same criteria was mentioned both in exclusion and inclusion criteria in the same paper, in most of the cases it was considered Justifiable (10 occurrences, 66 % of the cases).

The same analysis was performed for the categories mentioned as inclusion criteria in some papers and as exclusion criteria in others (but never as both in the same paper). In this analysis "Age" and "Adherence to protocol" were majorly described as inclusion criteria and in only one exception for each, described as exclusion, both cases were considered avoidable. The same situation happened for the category "incompatibility/compatibility with de procedure", although it was majorly an exclusion criteria instead of inclusion ("being able to tolerate the surgery"). For the remaining category "Response to previous treatment" three papers described it as inclusion criteria focusing on disease persistence and two papers as exclusion criteria one mentioning adverse effects and the other inefficacy of specific treatment to restore some physiological levels, both cases considered justifiable.

In the second phase of the study we investigated if the categories which were exclusively listed under inclusion and exclusion criteria could be found in the opposite criteria in at least one clinical trial. After the search the only criteria which continued exclusively as inclusion criteria were "Gender", and "Investigated condition" and as exclusion criteria "Conflict of interests/Ethical". It is worth mentioning that the Investigated condition varied across the studies and gender as well. The criteria "Generic nonspecific exclusion" was too unspecific to be searched.

Discussion

These data indicate: 1) although it is not possible (or maybe even desirable) to establish a list of definitive categories for selection criteria, they seem to be limited across study subjects, their identification could be interesting to promote a better comparison between studies and some efforts of this sort can be found in the literature (9). 2) there are only a few categories that could serve as a guide for inclusion or exclusion, since most of the categories (10 out of 18 or about 56%) were found either listed as inclusion or exclusion criteria or even as both in some cases; 3) in at least one third of the situations in which the same category was mentioned as inclusion and exclusion this could be avoided

and in most cases indicate a certain confusion about the role of the criteria for the study; 4) excluding particularities of cervical cancer studies only two categories remained as inclusion only and one as exclusion only.

Inclusion and exclusion criteria are not simply symmetric mirrorlike descriptions, inclusion should be generic, pointing to the desired characteristics of the population needed for the study, exclusion should specifically and clearly point to the characteristics of the subgroup of these participants which could not be accepted because they would not benefit or even be harmed by the procedures or would compromise the precise analysis of the results. So exclusion criteria, when used, should apply to the participants which would fit the inclusion criteria in first place, and should be very carefully planned to avoid potential risks or biases (both in favor or against the treatment) and patient harm.

The need of a clear differentiation between the functions of inclusion and exclusion criteria became evident only by the clear differentiation in the functions of efficacy and effectiveness trials: efficacy trials describe effects in selected study populations under experimental study conditions while effectiveness trials describe effects in NOT selected study populations under non-experimental (pragmatic) Real World Conditions (RWC). Efficacy studies are characterized by description of inclusion and exclusion criteria while effectiveness studies define only inclusion but no exclusion criteria. Efficacy and effectiveness trials have different goals, use different tools, and generate different messages (10–12).

Conclusion

Recommendations for inclusion and exclusion criteria should vary according to the purpose of the study, efficacy studies, which are designed to test the principle of the intervention, have to isolate any form of external contamination (either favoring or hindering particular results of the intervention) so the precise exclusion criteria should be carefully planned and justified. Effectiveness studies on the other hand deal with the effects of the intervention in a general population, they need observational designs instead of experimental and exclusion criteria basically do not apply. Considering every walk-in patient will receive treatment what is actually selected is the most suitable treatment and not the participant and comparisons occur across cohorts. It should be important to remark that even in these designs (usually termed pragmatic controlled trials) results cannot be simply generalized to any population, a thorough description of the key characteristic of the participants is essential to understand under which conditions it is safe to argue for the effectiveness of the intervention (i.e. comorbidities, gender, treatment association, and so on).

In summary, our results justify the following suggestions:

1. Before defining inclusion and exclusion criteria
 - a. The function of the study has to be defined;
 - b. The aim (or function) of a study may be to describe an experiment or an observation;

c. For description of an experiment both inclusion and exclusion criteria have to be defined. For description of an observation the definition of exclusion criteria is hard to justify and may even be counterproductive.

2. Inclusion criteria

a. have to be defined for all types of studies;

b. describe the conditions that are common to all subjects included in a study.

3. Exclusion criteria

a. are appropriate tools for experimental, i.e. explanatory studies that describe the efficacy; i.e. investigate the proof of a principle.

b. are not appropriate for observational studies i.e. which deals with the real world conditions (multiple risk factor and comorbidities) and treatment cannot be denied; i.e. investigates the treatment effectiveness.

c. describe the subgroup of the patients that fit the inclusion criteria, but present conditions that prevent them to take part in the study (e.g. ethical, biological, practical);

Declarations

Ethics approval and consent to participate

Not Applicable

Consent for publication

Not applicable

Availability of data and materials

Extraction tables and information of the analyzed paper are available under request to the authors.

Competing interests

The authors declare that they have no competing interests.

Funding

Not applicable

Authors' contributions

PCMM performed the main writing of the first version of the manuscript, the categorization and analysis of the results. CAG performed the paper search and selection, the planning of categories and categorization of the criteria. FP coordinated the activity, discussed the categorization and reviewed the

first version. All authors participated in the definition of the objectives, study design and discussion of the final version

Acknowledgements

Not applicable

References

1. Garritty C, Stevens A, Hamel C, Golfam M, Hutton B, Wolfe D. Knowledge Synthesis in Evidence-Based Medicine. *Seminars in Nuclear Medicine*. 2019;49(2):136–44.
2. Siontis KC, Ioannidis JPA. Replication, Duplication, and Waste in a Quarter Million Systematic Reviews and Meta-Analyses. *Circulation Cardiovascular quality and outcomes*. 2018;11(12):e005212.
3. Harris RG, Neale EP, Ferreira I. When poorly conducted systematic reviews and meta-analyses can mislead: a critical appraisal and update of systematic reviews and meta-analyses examining the effects of probiotics in the treatment of functional constipation in children. *American Journal of Clinical Nutrition*. 2019;110(1):177–95.
4. Cumpston M, Li T, Page MJ, Chandler J, Welch VA, Higgins JP, et al. Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. *The Cochrane database of systematic reviews*. 2019;10:ED000142.
5. Vandembroucke JP. Why do the results of randomised and observational studies differ? Statistical theory conflicts with empirical findings in several areas of research. *BMJ (Online)*. 2011;343(7834):1128.
6. Goldman L, Mudge GH, Cook EF. The changing “natural history” of symptomatic coronary artery disease: Basis versus bias. *The American Journal of Cardiology*. 1983;51(3):449–54.
7. Porzsolt Franz F, Galito N, Toledo-Arruda A, Thomaz T, Moraes C, Guerra T, et al. Efficacy and effectiveness trials have different goals, use different tools, and generate different messages. *Pragmatic and Observational Research*. 2015;47.
8. Porzsolt Franz F, Galito N, Toledo-Arruda A, Thomaz T, Moraes C, Guerra T, et al. Efficacy and effectiveness trials have different goals, use different tools, and generate different messages [Corrigendum]. *Pragmatic and Observational Research*. 2016;1.
9. Porzsolt F, Wiedemann F, Becker SI, Rhoads CJ. Inclusion and exclusion criteria and the problem of describing homogeneity of study populations in clinical trials. *BMJ Evidence-Based Medicine*. 2019;24(3):92–4.
10. Dubuisson N, Marta M, Gnanapavan S, Turner B, Baker D, Thomson A, et al. Inclusion criteria used in trials of people with progressive multiple sclerosis. *Multiple Sclerosis Journal*. 2018;1–5.
11. Luo H, Xu Y, Yue F, Zhang C, Chen C. Quality of inclusion criteria in the registered clinical trials of heart failure with preserved ejection fraction: Is it time for a change? *International Journal of*

Cardiology. 2018;254:210–4.

12. Ronconi JM, Shiner B, Watts B v. Inclusion and exclusion criteria in randomized controlled trials of psychotherapy for PTSD. *Journal of Psychiatric Practice*. 2014;20(1):25–37.
13. Porzsolt F, Weiss C, Weiss M, Jauch K, Kaplan RM. Perspective Article Demystification – a solution for assessment of real-world effectiveness? 2020;20(Appendix II):1–2.
14. Porzsolt F, Jauch K-W, Kaplan RM. 100 Years BAUHAUS Dessau/Germany: What designers of clinical studies can learn from designers and architects of BAUHAUS Dessau and hochschule für gestaltung ulm/Germany. *International Journal of Surgical Procedures*. 2020;3(1):133.

Tables

Table 1: Description of the identified eligibility criteria

<i>Adherence to protocol</i>	un/availability (exclusion/inclusion) and commitment to comply to the treatment protocol (e.g., commuting etc...)
<i>Age (range/minimum)</i>	age of the patient, either a minimum/maximum age (e.g. 18 years old) or some specific age range (e.g. between 18 and 70 years old).
<i>Autonomy</i>	possibility of the participant to understand the study protocol and characteristics, and voluntarily decide to participate (e.g psychological condition, language fluency)
<i>Conflict of interest/ethical</i>	any involvement of the patient with the study or research team which could jeopardize the study results. (e.g. sponsors, familial or personal relation).
<i>Described comorbidities</i>	specific diagnosis for the comorbidities present in the patient.
<i>Disease staging or characteristic</i>	staging of the disease, its severity or progression, which could either be required (inclusion) or jeopardize (exclusion) the study.
<i>Gender</i>	gender of the participant.
<i>Generic nonspecific exclusion</i>	when additional or subjective criteria different from the described was used to exclude the patient (e.g. paper #4 "Any condition that, in the opinion of the investigator, would interfere with evaluation of study treatment or interpretation of patient safety or study results")
<i>Incompatibility/Compatibility with the procedure</i>	any condition which might preclude the patient from undergoing any procedure used in the study (e.g. paper#2 "inability to enter the cervical OS"; paper #3 "Patients unable to undergo MRI for any reason"), or specific condition necessary for participating (e.g. paper#7 "the patients were able to tolerate the surgery")
<i>Informed consent</i>	signed consent of the participant to take part on the study.
<i>Life expectancy</i>	estimation of participant's lifetime.
<i>Medical History</i>	description of previous medical conditions related to the topics of interest of the study.
<i>Physiological functions</i>	specific physiological levels, organ function which may interfere (either positively or negatively) to the safety of the patient or the interpretation of the study results.
<i>Pregnancy/breastfeeding</i>	if the patient is pregnant, willing to be pregnant or breastfeeding.
<i>Investigated Condition</i>	identification of the primary disease studied.
<i>Response to previous treatment</i>	response to previous treatment.
<i>Standardized performance status</i>	accepted values of health status performance according to a standard scale (e.g. Karnofsky Performance Score; Eastern Cooperative Oncology Group performance status).
<i>Treatment</i>	any treatment (procedure, drug or trial), prior to the experiment or currently being taken or performed which might be necessary(inclusion) or interfere (exclusion) with the results or safety of the patient.

Table 2: Incidence of each category and type of criteria in which it was mentioned.

Criteria	Category	# of papers			
		In	Ex	In/Ex	Total
Inclusion only	Gender	3	0	0	3
	Informed consent	5	0	0	5
	Life expectancy	6	0	0	6
	Investigated condition	27	0	0	27
	Standardized performance status	26	0	0	26
Exclusion only	Autonomy	0	3	0	3
	Conflict of interest/ ethical	0	1	0	1
	Generic nonspecific exclusion.	0	1	0	1
Both	Adherence to protocol	3	1	0	4
	Age	12	1	0	13
	Described comorbidities	0	16	1	17
	Disease staging or characteristic	23	0	4	27
	Incompatibility/compatibility with the procedure	1	7	0	8
	Medical history	0	9	1	10
	Physiologic functions	17	2	2	21
	Pregnancy/breastfeeding	2	5	2	9
	Response to previous treatment	3	2	0	5
	Treatment	6	11	5	22

Table 3: Differences in selection criteria between efficacy (explanatory) and effectiveness (pragmatic trials).

	Efficacy trials (explanatory)	Effectiveness trials (pragmatic)
Inclusion criteria	Necessary and applied	Necessary and applied
Exclusion criteria	Necessary and applied	Do not exist

Figures

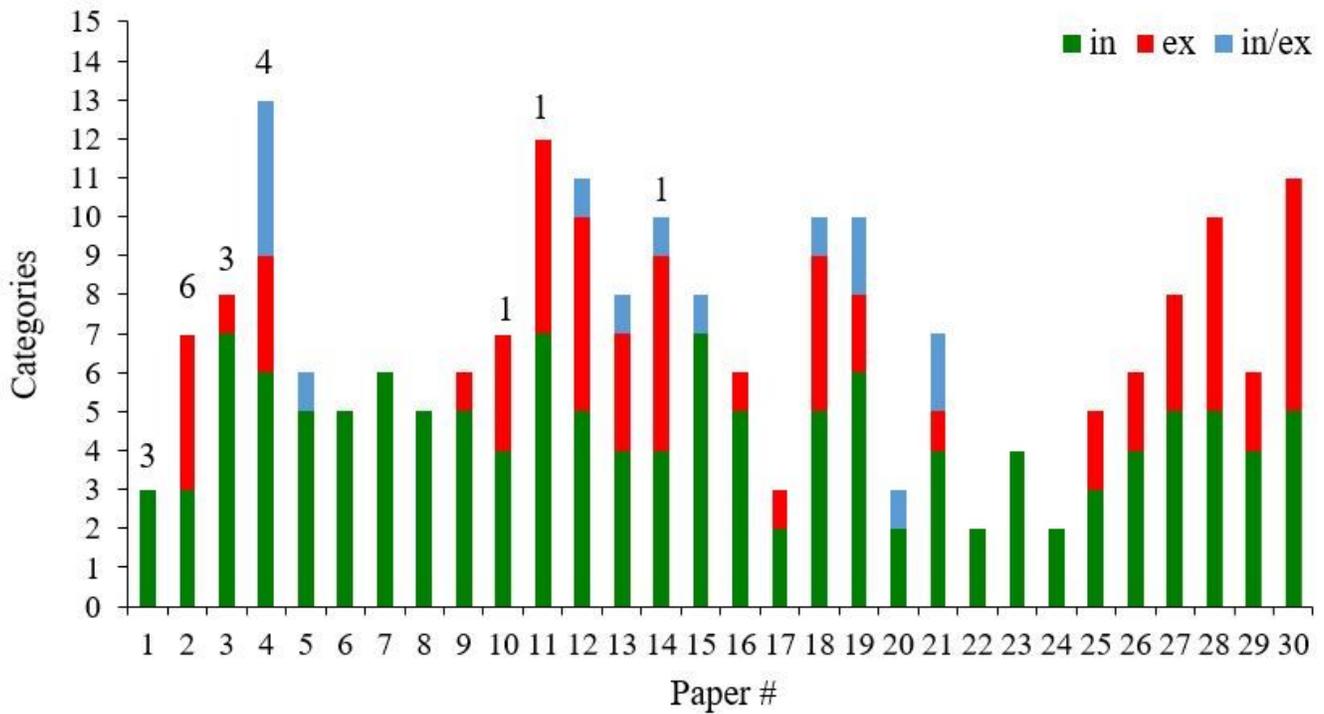


Figure 1

Total number of categories per paper distributed according to the criteria: inclusion (green), exclusion (red) and a third category of criteria that were uses as both inclusion and exclusion criteria in the same

paper (blue). The numbers above the bars indicate the number of categories added to the classification from the paper.

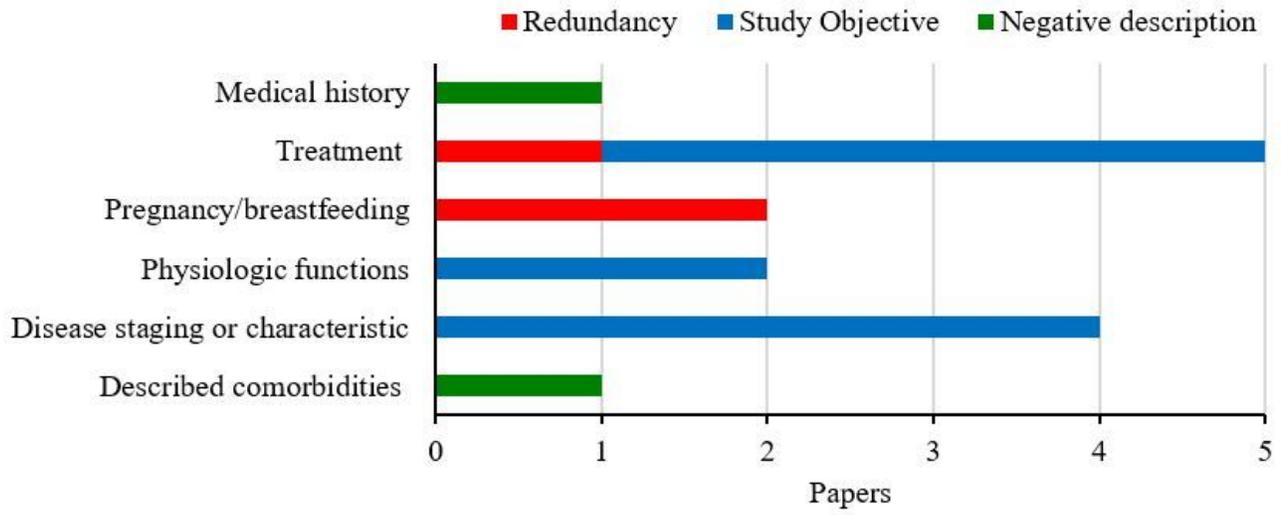


Figure 2

Incidence of the reasons the criteria was mentioned both in exclusion and inclusion criteria in the same papers.