Safety Science 158 (2023) 105998

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/safety

On the strength of arguments related to standardization in risk management regulations

Johan Ingvarson *, Henrik Hassel

Division of Risk Management and Societal Safety, Lund University, Box 118, 221 00 Lund, Sweden

ARTICLE INFO

Keywords:
Standardization
Risk management
Risk regulation
Argument analysis
Scoping study
Risk standard
Risk management regulation

ABSTRACT

Risk-informed decision-making is a key concept in regulatory regimes covering a wide range of technical disciplines and risk domains. Increased standardization in terms of highly prescriptive and detailed requirements is observed in various risk management regulations and application areas. This development has spurred a discussion with equally strong voices promoting and opposing standardization of risk. The aim of this paper is to explore the knowledge base from scientific literature related to effects of standardization of risk and to assess the strength of the arguments used in the discussion. Through a scoping study and a systematic argument analysis, this study contributes to the debate on the most adequate approach to standardization in risk management regulations. Conclusions from the study are that effects of standardization of risk management regulations are not widely covered in scientific literature and empirical evidence is largely lacking to support the arguments used. The argument analysis indicate that many arguments are assessed as being weak, interconnected and even contradictory. It seems not possible to come to a conclusion and reach agreement in the discussion if the strengths of standardization outweigh the weaknesses. Thus, the discussion should be in the form of dialectic rather than debate. Acknowledging, understanding and addressing the various arguments in favour of and rejecting standardization, as well as their respective context, is essential when designing risk management regulations. Future research efforts are required to further explore when and how the risk management and risk governance processes can be standardized and the appropriate level of standardization.

1. Introduction

Managing and governing risk is at the core of regulatory regimes covering land-use planning, process safety, occupational health and safety, security, environmental protection and many other technical disciplines and risk domains. The practices of risk-informed decision-making processes are implemented differently in different regulations. Some regulatory regimes require risk assessments to be performed and used for decision-making without defining how they should be conducted and used. Others are more standardized in terms of providing prescriptive requirements on key risk management components such as risk analysis methodologies, assumptions, modelling tools, data input, evaluation criteria, etc. Viewing the level of detail prescribed in risk management regulations as representing the level of standardization, highly prescriptive regulatory regimes then represent high level of standardization whereas regulations with lesser prescriptive requirements represent low level of standardization. Standardization in risk management regulations can be applied to individual elements of the risk management and risk governance processes (e.g. risk analysis) or the entire process.

Standardization and the use of standards in risk management regulations related to land-use planning is increasingly popular (Ingvarson, 2020, p. 395). The development towards increased standardization is also observed in diverse risk domains such as terrorism and security risk management (Jore, 2020), cybersecurity (Skotnes, 2020) and disaster risk management (Morsut, 2020). Increased standardization, irrespective of risk domain, is also met by critical voices that problematize the value of standardization of risk management regulations (Aven and Ylönen, 2019; Olsen, 2020a). Ingvarson (2020) indicates that analyses of the effect or value of applying risk management processes with different levels and types of standardization are not readily available in research literature. It is suggested that positioning as proponents or opponents of standardization of risk management regulations is based on mostly practical considerations and intellectual reasoning rather than empirical data on the effects of standardization. Arguments from the ongoing debate on whether standardization of risk management

* Corresponding author.
E-mail address: johan.ingvarson@risk.lth.se (J. Ingvarson).

https://doi.org/10.1016/j.ssci.2022.105998
Received 24 January 2022; Received in revised form 14 October 2022; Accepted 27 October 2022
Available online 11 November 2022
0925-7535/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

regulations is desirable represent both the strengths and weaknesses of standardization (Ingvarson, 2021). But how strong are the arguments supporting or rejecting standardization of risk management regulations? If empirical evidence, typically representing rather strong scientific evidence, is missing, the strength of arguments is increasingly critical when evaluating or justifying the level of standardization in risk management regulations.

The aim of the present paper is two-fold; first, the paper aims at adding to the current knowledge base by presenting an overview of studies of the effects of standardization and the use of standards in risk management regulations found in scientific literature. Secondly, the aim is to assess the strength of the arguments promoting or opposing standardization of risk management regulation found in the literature review. Risk management regulations in this context refer to policies and governing documents (international and national laws and legislation, rules, recommendations, guidance, etc.) aimed at governing the risk management process rather than the development or manufacturing of products, components, etc. Through a systematic argument analysis, contribution is made to the debate on the most adequate approach to standardization in risk management regulations. The methods used for the systematic literature review using a scoping study approach and the argument analysis are presented in section 2. Details of how the scoping study was performed and its results are covered in section 3. The findings from the scoping study results in terms of arguments related to standardization in risk management regulations are presented in section 4 and analysed in section 5. In section 6, the scoping study and argument analysis results are discussed and some conclusions are presented in section 7.

2. Method

The approach in this paper is based on an initial systematic review of scientific literature to explore the effects of standardization and the use of standards in risk management regulations and to identify arguments related to such standardization. Following the advice of Munn et al. (2018), a literature review using a scoping study approach forms the basis of this paper. This is presented in section 2.1. The results of the scoping study in terms of arguments in favour of or rejecting standardization in risk management regulations, are then analysed. The approach to assessing the strength of arguments is presented in section 2.2. The approach has been designed specifically for this paper and is based on traditional concepts for argument analysis and inspired by risk science contributions related to validity of risk assessments and strength of knowledge classification.

2.1. Scoping study

Acknowledging that a literature review "requires more than just compiling a collection of references about a topic" (Kowalczyk and Truluck, 2013), structure, thoroughness, methodological rigour and transparency are key traits. For this purpose, a wide array of methods and approaches to literature reviews is available. Considering the overlap of methodologies and the lack of "an internationally agreed set of discrete, coherent and mutually exclusive review types" (Grant and Booth, 2009), selecting the most appropriate method for literature review is a delicate task. To support the selection process of an adequate approach and methodology for literature review, descriptions and comparisons of different review types have been consulted, such as Ferrari (2015); Poth and Ross (2009); Randolph (2009); Tranfield et al. (2003).

Guided also by the well-recognized taxonomy of literature reviews developed by Cooper (1985), a scoping study approach is selected for this literature review. According to Mayset al. (2001, p. 194, cited in Arksey and O'Malley, 2005), a scoping study "can be undertaken as stand-alone projects in their own right, especially where an area is complex or has not been reviewed comprehensively before". This definition of a scoping study fits well with the aims of the current paper.

Arksey and O'Malley (2005) identified four common reasons for undertaking a scoping study:

1. To examine the extent, range and nature of research activity.
2. To determine the value of undertaking a full systematic review.
3. To summarize and disseminate research findings.
4. To identify research gaps in the existing literature.

With respect to the need for a comprehensive overview of standardization in risk management regulations in scientific literature, at least three of the four reasons are applicable. The value of systematic review (item 2 above), with its intrinsic differentiation of studies of good and poor quality (Poth and Ross, 2009), is not a primary concern of this paper as the purpose is rather to chart the entire landscape of arguments for/against standardization of risk management regulations.

The methodological framework of Arksey and O'Malley (2005) and the subsequent enhancements of the framework by Levac et al. (2010) and Daudt et al. (2013) is applied in this paper. The framework is well-designed to meet the ambitions of the current scoping study and the approach is also commonly used and advocated (Colquhoun et al., 2014).

The Arksey and O'Malley framework for scoping studies includes five stages and an optional sixth step of "Consultation", which have been included in the present paper. Each of the steps and how it is applied in this scoping study is described in detail in section 3.

2.2. Argument analysis

An important aspect of analysing an argument is the strength of evidence supporting the argument. An argument underpinned by relevant and undisputed evidence is typically interpreted as strong. Similarly, arguments based on hearsay or personal beliefs are seemingly weaker. The relative strength of evidence is often ranked using an evidence hierarchy or levels of evidence. With this approach, the rigour of the evidence is evaluated based on degree of evidential support. No matter if the hierarchy of evidence is illustrated by the five or seven levels of evidence common in scientific literature (e.g. Nutley et al., 2013), strength of evidence is primarily related to the quality and quantity of supporting empirical data. Hence, using strength of evidence as a dimension of argument strength requires access to an adequate pool of empirical studies. Findings from the scoping study indicate that there are few studies including empirical data related to effects of standardization of risk management regulations (Ingvarson, 2021). Considering the unavailability of solid empirical bases for arguments found in the scoping study, it was decided not to include hierarchy of evidence as a method for argument analysis in the present paper.

In search for an appropriate method for argument analysis, the roots of argumentation theory were required. Arguments as the basis for persuasion have been at the heart of intellectual debates and discussions since the dawn of science. Theory of argumentation has evolved into a rich interdisciplinary area of research encompassing perspectives from philosophy, psychology, linguistics, communication science, and more. Founded in ancient Greece, Aristotelian rhetorical theory differentiates three means of which persuasion may occur; through a speaker's character (*ethos*), an audience's emotional state (*pathos*), and the argument itself (*logos*) (Russell and Reimer, 2018). Specifically focusing on *logos*, a number of approaches and tools have been developed over the years to support the art of argumentation, i.e. concluding through logical reasoning with the use of arguments. Landmark contributions to the use of arguments include Toulmin's (2003 [1958]) model of six interrelated components for analysing arguments (claim, data/grounds, warrant, backing, qualifier, and rebuttal), Johnson and Blair (2006 [1977]) ARS (Acceptability, Relevance, and Sufficiency) criteria for a logically good argument, and the use of argumentation schemes representing structures of inference for common types of arguments (see for example

Table 1
Criteria for classifying argument strength in terms of soundness and relevance.

Dimension	Label	Criteria
Soundness	Strong	All desired qualities for the soundness dimension are assessed as Clearly fulfilled.
	Moderate	Less than three desired qualities for the soundness dimension are assessed as Clearly fulfilled AND no desired qualities for the soundness dimension are assessed as Clearly not fulfilled.
	Weak	One or more of the desired qualities for the soundness dimension are assessed as Clearly not fulfilled.
Relevance	Strong	All desired qualities for the relevance dimension are assessed as Clearly fulfilled
	Moderate	Less than three desired qualities for the relevance dimension are assessed as Clearly fulfilled AND no desired qualities for the relevance dimension are assessed as Clearly not fulfilled.
	Weak	One or more of the desired qualities for the relevance dimension are assessed as Clearly not fulfilled.

Walton et al. (2008) for a collection of 65 argumentation schemes).

Critically analysing arguments aims at distinguishing between strong and weak (or even fallacious) arguments. A common denominator for the various approaches to argument analysis is the essential judgement of the argument's *soundness* and *relevance*. Soundness relates to how the argument contributes to increasing the acceptability of a standpoint (van Eemeren and Snoeck Henkemans, 2017). A sound argument builds trust in the support of a claim by use of validity, i.e. whether a particular conclusion inevitably follows from its premises, and confidence in the premises being true. In contrast, an unsound argument does not contribute to the acceptability of the standpoint either because it is invalid or because it is valid but its premises are untrue (Chatfield, 2018).

A perfectly sound argument, i.e. an argument that is both valid and based on premises that are justifiably true, can still be considered weak in case it has limited or no bearing on the applicable claim. A strong argument also has to be contextually *relevant*; it has to be closely relating to the subject or point at issue. The challenge with arguments of low relevance is that not only do they fail to prove or cast doubt on the concerned issue, they are distracting and prevent a dialogue from progressing toward resolution of the issue (Walton, 2006).

An argument's strength, corresponding to argument quality, is in the present paper determined by the two dimensions of soundness and relevance. The more sound and relevant the argument is, the stronger it is regarded. Consequently, arguments assessed as being weaker are regarded less sound and/or relevant.

Argument analysis is inevitably subjective as it is largely built on the assessor's knowledge and value preferences. Both soundness and relevance of an argument are therefore preferably measured using a multi-step scale rather than on a binary scale, i.e. sound/unsound or relevant/irrelevant, to capture the complexity and impreciseness of value-laden assessment of argument strength. Even when using a non-binary scale, the fundamental problem that any assessor of arguments will encounter is limited knowledge and information. Whether trying to determine the acceptability of a standpoint, the validity of arguments, or the justification of premises, in all these situations the assessor's judgements and interpretations are often limited by the epistemic situation of not knowing everything that would be required to evaluate an argument's quality objectively (Hoffmann, 2018). Additional information that is not conveyed to the assessor, that the assessor fails to recognize or that the assessor interprets in a way that is not in line with the best available science, can significantly alter the assessment. The assessment of strength of the arguments performed in this paper should therefore be seen as possible to dispute by providing new knowledge or challenge the interpretations of existing information. The response to this dilemma in this paper is to carry out an as transparent assessment as possible. This is done by developing a set of criteria for argument strength and to include appendix B with a summary of all assessments made. But even with seemingly objective criteria, argument analysis remains subjective and debatable.

In search of argument criteria, both the theoretical grounds for argument quality and how argument quality can be operationalized through receiver evaluations should be considered (see for example Russell and Reimer (2018) for an overview of argument quality criteria). Since the meaning of *argument* varies in different traditions (Hoffmann, 2018), it is necessary to develop quality criteria relevant for the discourse of the discipline in which the argument analysis is performed.

Risk assessment is a cornerstone in risk management and as such the validity, quality and trustworthiness of risk assessments are topics of concern in risk research. Over the years, considerable efforts have been used to develop approaches to assess validity/quality/trustworthiness in risk assessments, primarily quantitative risk assessments (QRA), and risk assessment processes (see for example Goerlandt et al. (2017) for an overview). Theoretical, methodological and empirical contributions in risk science have been used as sources of inspiration when developing criteria for assessing argument strength in this paper. Particularly the three-level (unreliable; reliable but invalid; reliable and valid) classification-based method for trustworthiness assessment by Zeng and Zio (2017) and the four-level (unrepeatable; invalid; valid but inaccurate; accurate but challengeable) maturity model for QRA quality assurance and review by Rae et al. (2014) have been used as input.

In risk research, the "strength of knowledge" concept is used to characterize the background knowledge in risk assessment. For this purpose, a three-tiered set of criteria for making judgements of the core risk component of uncertainty is developed (Askeland et al., 2017; Flage and Aven, 2009). Askeland et al. (2017) refer to and label the strength of knowledge as either "strong", "moderate" or "weak", a structure and vocabulary that is used as input when developing criteria for argument analysis in the present paper. The final phrasing of labels for argument strength also considers important aspects of wording, as discussed for example in Flage and Aven (2017), Goerlandt and Reniers (2017) and Dent et al. (2010), Flage and Aven (2018), to prevent ambiguity and unclear delineation when defining criteria.

The strength of an argument, in terms of the two dimensions soundness and relevance, is here assessed based on to what extent a specific set of desirable qualities are achieved. The soundness dimension includes the following desirable qualities:

1. A conclusion inevitably follows from its premises.
2. There are strong reasons for believing the premises are true.
3. Argument is supported by evidence in the form of empirical source material/data OR non-empirical source material/data of high quality.

A deliberate and important distinction is made between the value of empirical versus non-empirical source material/data. An argument is considered more sound if supported by empirical evidence than if only non-empirical evidence is presented. This means that an argument must be supported by non-empirical evidence of high quality (e.g. thorough logical inference, comprehensive intellectual reasoning and reflection, etc.) to balance absence of empirical evidence in order to fulfil quality 3 above.

Furthermore, the relevance dimension includes the following desirable qualities:

A. Premises are contextually connected with the subject or point at issue.
B. There is a pertinent substantial relation between the premises and the conclusion of the argument.
C. Evidence presented have bearing on the truth of the claim at issue.

Each of the six desirable qualities are assessed for each argument as; a) Clearly fulfilled; b) Clearly not fulfilled; or c) Neither clearly fulfilled

nor clearly not fulfilled. Option c) is included as there could be instances when it cannot be clearly determined whether the quality is fulfilled or not, e.g. where additional data collection is required to be able to make a clear-cut assessment. Based on the assessments, argument soundness and relevance are then labelled as *Strong*, *Moderate* or *Weak* in accordance with the criteria specified in Table 1. Since the intention is not to perform a meta study (i.e. assessing the synthesis of all evidence related to one argument from all publications), if the same argument is assessed differently on the desirable qualities in different publications from the scoping study, the highest assessment is chosen.

3. Conducting the scoping study

Each of the stages 1 to 5 of the Arksey and O'Malley methodological framework are explored in detail in section 3.1 to 3.5. The optional consultation exercise is included in Stage 3 Study selection.

3.1. Stage 1: Identifying the research question

The starting point is to identify a research question on which the search strategy is built. Arksey and O'Malley (2005) suggest to maintain a wide scope when phrasing the research question to include a breadth of coverage. To ensure that the research question also provides the direction and focus needed to inform subsequent stages of the research process, Levac et al. (2010) recommend that a broad research question is combined with a clearly articulated scope of inquiry. For clarity, it is also recommended to mutually consider the purpose of the scoping study with the research question.

The research question is defined as: *How are different ways of standardizing risk management regulations described in the scientific literature and have the effects of these been evaluated?* The phrasing of the research question is intentionally wide by including different types of standardization and different kinds of risk applications (i.e. from various risk domains and disciplines).

3.2. Stage 2: Identifying relevant studies

To answer the research question, Arksey and O'Malley (2005) suggests searching for research evidence using electronic databases, reference lists, hand-searching of key journals, and using existing networks, relevant organizations and conferences. The primary type of information source for this scoping study is electronic databases (see section 3.2.1 for details). The Scopus (owned by Elsevier), Web of Science (Clarivate Analytics) and Academic Search Complete (EBSCO) databases were selected as they are multi-disciplinary, cover a wide range of research fields, and are owned by various publishers. In line with the research question relating to scientific literature, no grey literature was searched.

Following the completion of database searches, a hand-search of the International Journal of Standardization Research (IJSR) and its predecessor International Journal of IT Standards and Standardization Research (IJITSR) was conducted (see section 3.2.2). To further build confidence that all relevant studies were identified, the reference lists from the most relevant literature were scanned to identify other papers that may not have been found in the initial search (section 3.2.3). Finally, recognized scholars from the standardization and risk disciplines were consulted to reassure that the most distinguished papers were identified in the scoping study as described in section 3.2.4.

The scoping study was performed by the authors with the support of a university librarian. The risk of inconsistency was mitigated by a work process allowing for regular meetings for calibration and experience and information transfer and the use of clear inclusion/exclusion criteria. The obvious risk of omitting essential literature was reduced by performing the optional consultation exercise as part of Stage 3: Study selection.

3.2.1. Database search

The search query used in the database queries was based on a Boolean approach in combination with an iterative work process. Starting with the distinctive keywords *standardization* and *risk* from the research question, a relevant search query was fine-tuned using a set of key words and related terms and synonyms. The final search strategy used was defined as "(standard*) AND ("risk management" OR "risk analysis" OR "risk assessment" OR "QRA" OR "risk governance" OR "risk criteria" OR "risk communication" OR "risk acceptance" OR "risk tolera*") AND (effect OR result* OR consequence* OR value OR outcome OR impact OR benefit) AND (evaluat* OR assessment* OR review* OR analys?s OR rating* OR appraisal*". By using proximity operators and restricting the search to only peer-reviewed publications in English labelled in the respective databases with relevant subject areas or research category, the initial overwhelming number of records of many tens of thousands was reduced to a more manageable 1 622 records.

3.2.2. Hand-searching of key journals

Hand-searching of key journals is labour-intensive so finding an appropriate balance between the resources used and the output gained is critical. Using the internationally recognised databases, most prominent and relevant journals are already included in the data underpinning the search query results. When assessing the results from the database search, it was recognized that the key journals typically used in risk research are represented among the records found (e.g. Safety Science, Journal of Risk Research, Reliability Engineering and System Safety, Journal of Safety Research, Journal of Hazardous Materials, and Risk Analysis), implying that these sources of information are fully covered using the database search. Hence, the hand-searching of journals was focused on journals that publish research related to standards and standardization but are not represented in the database search results. The International Journal of Standardization Research (IJSR) and its predecessor International Journal of IT Standards and Standardization Research (IJITSR) were selected as these journals publish research related to standards and standardization in different types of organizations and across industries. As such, these journals were considered a potentially valuable reference source for empirical research and real-world case studies and examples. Hand-searching IJSR and IJITSR resulted in 138 articles feeding into the title analysis of stage 3: Study selection.

Considering that standardization bodies are key players in standardization and the considerable investments they have made over the years in convincing other stakeholders of the value of standardization, turning to these organization to search for input to this scoping study may result in valuable sources of information. ISO, for example, has compiled research from academia, international organizations and ISO Members and more. The ISO Research Library contains a collection of publications about standards, standards development and the impacts of standardization. The collection is easily accessible via the ISO web site (library.iso.org) and contains both academic publications and grey literature. However, since the scientific literature available through ISO Research Library is not exclusively found in their collection (rather, publications are made available via links to their respective original source), it is assumed that the publications relevant to this scoping study are also identified using the database searches and hand-searching of key journals. This assumption was verified by a sample of a dozen journal articles found in the ISO Research Library that were also retrieved using the same databases as in the database search in section 3.2.1. Database search.

3.2.3. Reference lists review

Arksey and O'Malley (2005) point out the value of checking the bibliographies of studies found through the database searches to ensure they had been included in the scoping exercise. Relevant references found when reviewing the bibliographies of the most relevant and authoritative studies (so called "pearl papers") from the database search and consultation exercise were added to the list for study selection.

A sort of reference list review that was not included in Arksey's and O'Malley's original methodological framework, is citation search. The citation search reveals relevant publications by looking at what a

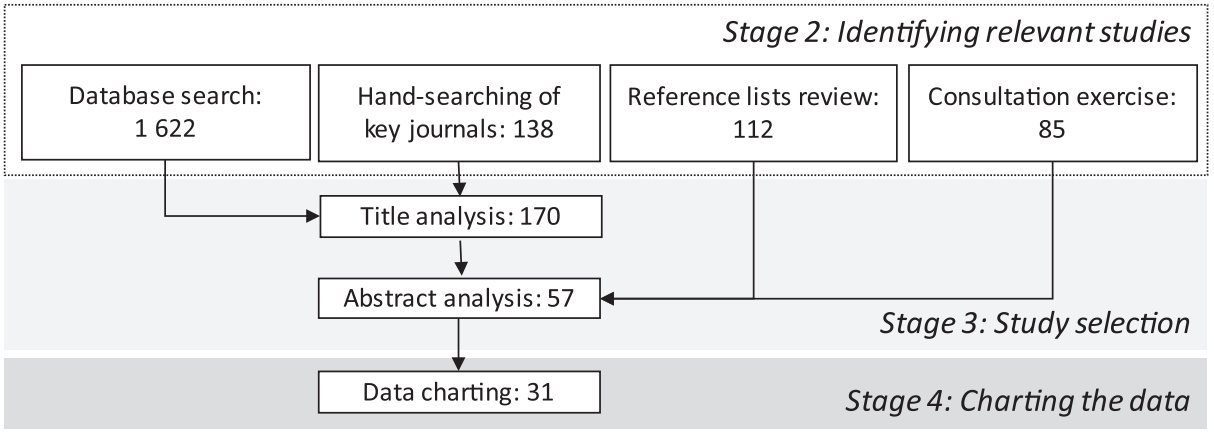

Fig. 1. Overview of stages 2 to 4 of the application of Arksey's and O'Malley's methodological framework for scoping studies.

publication has referenced and who has used that publication as a reference at a later stage. Some of these related articles may be of relevance for the specific research question of this scoping study. With this tool, relevant contributions from unexpected disciplines may turn up. Also, citation search allows for following up on how discussion, arguments and ideas are developed. Citation searching is facilitated by database features made available by the database providers. In this scoping study, Scopus was used for citation search.

Reviewing the reference lists and performing citation searches of relevant literature help construct a web of knowledge of the defined topic. This added a total of 112 publications to stage 3: Study selection.

3.2.4. Consultation exercise

The final step of identifying relevant studies was to consult recognized experts in the related disciplines. The intention of the consultation exercise is to add to the methodological rigour and to provide additional sources of information. In this scoping study, 30 internationally recognized practitioners and scholars were contacted by email. Each expert was tasked with an open question of listing the most distinguished contributions to research related to standardization of risk management regulations. In some cases, additional knowledge and experience on the topic was transferred via video meetings.

The consulted experts were selected based on their respective contribution to research in related disciplines. An intentional leverage towards experts from academia with a primary focus on the standardization perspective was used to balance the authors' main orientation in the risk domain. Experts were selected using a snowballing approach with initial contacts made to leading authors of publications found in the database search and representatives from standardization organizations such as the International Organization for Standardization (ISO), International Telecommunication Union (ITU), European Telecommunications Standards Institute (ETSI), and European Academy for Standardisation (EURAS). The 22 respondents included 16 men and 6 women affiliated with research institutions (19) and standardization organizations (3) across Europe (19) and North America (3).

Based on input from the respondents, another 85 publications were added to the abstract analysis of stage 3: Study selection. Since the experts were asked to contribute with contributions to research related to standardization of risk management regulations in general and not specifically to the effects of standardization, the majority of the suggested publications from the consultation exercise were not identified through the other data sources.

3.3. Stage 3: Study selection

The search strategy used when identifying relevant studies (Stage 2) resulted in a large number of studies that are irrelevant when addressing the research question. This is expected when seeking for breadth rather than depth. The traditional response to this in systematic review methods is to develop a set of inclusion and exclusion criteria based on the research question. Inclusion criteria are characteristics that the prospective publications must have if they are to be included in the study. Exclusion criteria are those characteristics that disqualify prospective publications from inclusion in the study. The use of inclusion/exclusion criteria is introduced to ensure consistency in decision-making in study selection. Following the suggested approach of Arksey and O'Malley (2005), criteria were devised *post hoc* based on increasing familiarity with the literature and then applied to all publications identified in Stage 2. The applied set of inclusion and exclusion criteria and its rationale is presented in Appendix A.

The first step of selecting relevant publications for further study was to perform a title analysis of the 1 760 publications identified in Stage 2 (1 622 from database searches and another 138 from hand searching of key journals). By assessing the title against the inclusion/exclusion criteria, publications that are irrelevant to the research question were removed. To ensure that the title analysis did not result in the rejection of relevant publications, borderline cases were retained for the next step of abstract analysis. Although not part of Arksey's and O'Malley's original methodological framework, the title analysis was necessary to handle the substantial output from the database and hand-searching in Stage 2. Following the recommendation by Levac et al. (2010), the title analysis was performed by two researchers each independently reviewing the titles yielded from the search strategy for study selection. In line with the approach of not rejecting border line cases, if one of the researchers selected a publication for further analysis, the publication was retained. The title analysis resulted in 170 publications being selected for abstract analysis. Of the 170 publications, about 60% were identified by both researchers and 40% by one of the researchers.

The publications selected in the title analysis were combined with the 112 publications from the reference list review and the 85 publications from the consultation exercise as input for abstract analysis. Unfortunately, the abstracts of 14 publications were not available through the library databases or the internet so they had to be removed resulting in a total of 353 publications for abstract analysis. In the abstract analysis, the same inclusion/exclusion criteria were used for selecting the publications to be included in the study basis. Copies of the full article were obtained for those publications that appeared to fit with the research question. A similar approach was used as for the title analysis where all abstracts were reviewed by two researchers. To qualify for inclusion in the study basis, both researchers had to agree on the study selection. If the relevance of a study was unclear from the abstract, border line cases were included in the study basis. The final number of publications in the study basis feeding into the data chartering of Stage 4 was 57. An overview of Stage 3: Study selection is provided in Fig. 1.

Table 2
List of pro (P) and con (C) arguments commonly found in the scoping study.

ID	Argument	Brief description	Argument focus	Source publication
P1	Contribute to uniformity	Contribute to uniformity (e.g. defining, identifying, analysing, assessing and communicating risk) making risk more comparable and therefore more manageable.	Standard	Aven and Ylönen (2019) Olsen (2020b) Kica and Bowman (2012) Antonsen, et al. (2012) Aas and Johnsen (2007) Jore and Moen (2015)
P2	Increased likelihood of identifying relevant risks	Support the identification and inclusion of all relevant risks, including those not of particular interest to the party performing the analysis.	Standard	Aven and Ylönen (2019) Aas and Johnsen (2007) do Nascimento, et al. (2020)
P3	Bench-marking opportunities	Compare practices with internationally recognised benchmarks and allows a common approach to be taken across boundaries (i.e. a single international reference point).	Standard	Aven and Ylönen (2019) MacRae (2011)
P4	Increased safety	Increase the safety of products and safety for humans and the environment.	Standard	Aven and Ylönen (2019) do Nascimento, et al. (2020) Murakami (2016) MacRae (2011)
P5	Provide cost-effective solutions	Using standards, there is a potential for lower costs when there is no need to demonstrate compliance or explore alternative solutions.	Standard	Nyvik, et al. (2021) Belluck, et al. (2006)
P6	Regularly updated	Standards are regularly reviewed and updated providing opportunities to introduce recent developments.	Standard	Aven and Ylönen (2019) Kica and Bowman (2012)
P7	Offers predictable low risk levels	Provide lower risk levels by offering predictability (i.e. if a standard is complied with, a predicted adequate risk level is achieved).	Standard	Thompson (2020) do Nascimento, et al. (2020)
P8	Decreased administrative burden	Provide predictable and predefined targets, approaches and procedures to which the organization can streamline internal work processes.	Standard	Kica and Bowman (2012) Belluck, et al. (2006) MacRae (2011)
P9	Effective allocation and use of resources and expertise	Prevents “re-inventing the wheel” allowing organizations to capitalize on the expertise/resources of others rather than requesting all organizations to have all expertise/resources available in-house.	Standardization	Aven and Ylönen (2019) Foliente (2000) Lindøe and Baram (2020)
P10	Best available knowledge	Expert knowledge is stored in rules and technical solutions representing “best available knowledge”.	Standardization	Aven and Ylönen (2019) Nyvik, et al (2021) Skotnes, 2020 Lindøe and Baram (2020) Antonsen, et al (2012)
P11	Compliance verification	Standards are related to expectations of governance and represents defendable, auditable processes and a way to demonstrate compliance. Standards are easy for third party to check and regulators to enforce.	Standardization	Skotnes, 2020 Foliente (2000) Aven and Ylönen (2019) Belluck, et al. (2006)Jore and Moen (2015)
P12	Facilitate risk governance	Help facilitate the integration of several sources of information in multi-actor risk assessment processes and reduce flexibility in risk governance.	Standardization	Tehler, et al (2020) Olsen (2020a) Aas and Johnsen (2007)
P13	Hegemonic approach to managing risk	Contribute to building a legitimate hegemonic approach to the management of risk that is especially helpful for topic areas that are less often recognized and where organizations have little experience.	Standardization	Jore, 2020 Aas and Johnsen (2007)
C1	Produce conformity and homogeneity	Standardization produces conformity and homogeneity between stakeholders, while risk management and governance requires diversity and heterogeneity.	Standard	Aven and Ylönen (2019) Olsen (2020a) Jore, 2020Jørgensen and Lindøe (2020)
C2	Compliance-driven	Specifies minimum solutions and does not encourage higher ambition levels. Compliance with the standard(s) is in focus, not the optimal level of risk.	Standard	Nyvik, et al (2021) Clark-Ginsberg and Slayton (2019) Jore and Moen (2015)
C3	Fails to keep up with changes	The standard fails to keep up with changes and becomes outdated due to new knowledge or new alternative solutions becoming available/ developed.	Standard	Nyvik, et al (2021) Clark-Ginsberg and Slayton (2019) Aas and Johnsen (2007)Jore and Moen (2015)
C4	Inflexible and costly solutions	The standard may not be developed for the context specific conditions and appear less flexible which may result in less effective risk-reducing measures and costly solutions.	Standard	Nyvik, et al (2021) Raj and Lemoff (2009) Foliente (2000) Jore and Moen (2015)
C5	Universally agreed norms not available	Normative judgements are at the core of both risk and standardisation and since there are no/few universally agreed set of norms, standardisation is challenging. This also includes the discussion on what level of risk is acceptable and to whom?	Standard	Juhl, 2020 MacRae (2011)
C6	Increased bureaucracy	Standards may include administrative requirements that are cumbersome to some organizations (typically smaller organizations with less comprehensive management systems).	Standard	Antonsen, et al. (2012) MacRae (2011) Aas and Johnsen (2007)
C7	Barrier to innovation	Barrier to innovation by locking the industry into compliance with ‘one-size-fits-all’ rules, thereby disincentivizing and limiting flexibility and adaption to new technologies.	Standardization	Aven and Ylönen (2019) Foliente (2000) Lindøe and Baram (2020) Onofrio, et al. (2015)
C8	Inadequate transparency	Challenges in terms of legitimacy due to inadequate transparency (i.e. public access to information and decision-making procedures).	Standardization	Kica and Bowman (2012) Lindøe and Baram (2020)
C9	Inadequate stakeholder representation	Challenges in terms of legitimacy due to inadequate or imbalanced stakeholder representation (e.g. involvement of experts and presence and participation of unions, workers, the public, other industries, etc.).	Standardization	Aven and Ylönen (2019) Kica and Bowman (2012) Lindøe and Baram (2020) Urciuoli (2018)
C10	Does not address complex and emerging risks	Complex risks are associated with limited historical data (i.e. large uncertainties) which is a typically required for successful standardization. This may also reduce comparability with other complex risks and could reduce the capacity to discover and understand new risks on the horizon.	Standardization	Jore, 2020 Olsen (2020a)
C11	May generate false sense of safety	Standardization may reduce our attention to unthinkable threats and risks it is easy to talk about but difficult to handle, potentially generating a false sense of safety/security.	Standardization	Skotnes, 2020 Olsen (2020a) Jore, 2020 Clark-Ginsberg and Slayton (2019)
C12	Deterioration of technical competence	Reliance on industry standards and practices may lead to deterioration of own technical competence and local knowledge.	Standardization	Lindøe and Baram (2020) Antonsen, et al. (2012)
C13	Reduced ownership and quality	Standardization may lead to a loss of ownership and reduced motivation to perform risk analyses (or other risk management activities) and thus reduce the quality of assessments/activities.	Standardization	Tehler, et al. (2020) Olsen (2020a) Antonsen, et al. (2012)

3.4. Stage 4: Charting the data

The next stage of the work included extracting (or "charting") the data from the study basis. The extraction of data "should be a logical and descriptive summary of the results that align to the objective/s and question/s of the review" (Peters et al., 2015, p. 144). To organize data, Arksey and O'Malley recommend a "data charting form" to sort data according to key themes and issues. Accordingly, a data charting form was developed to facilitate data extraction while creating a uniform basis for analysis. As suggested by Levac et al. (2010), charting the data was considered an iterative process in which the data charting form was continually updated. A preliminary set of categories was developed to reflect the research question and to accommodate both quantitative and qualitative data analysis. Following the data extraction of a sample of publications from the study basis, some categories were fine-tuned and sub-categories were included.

The first step of charting the data was to validate if the publications selected for the study basis relate to the research question. This was required since publications could be interpreted as being relevant to the research question when assessing the title and/or abstract, but when reviewing the full paper, a more accurate interpretation can be made of the publication content, possibly resulting in exclusion from the study. The validation was done by re-assessing the inclusion/exclusion criteria in Appendix A following full text review of each of the 57 publications in the study basis. 26 publications were excluded following the re-assessment of the inclusion/exclusion criteria. Out of the 31 publications selected and used as input when populating the data charting form, 14 were identified through the database search. The additional 17 publications were identified using the consultation exercise.

3.5. Stage 5: Collating, summarizing and reporting the results

As pointed out by Arksey and O'Malley (2005), a scoping study aims at presenting a narrative account of existing literature; it does not – as is the case with systematic reviews or *meta*-analysis – primarily seek to synthesize evidence or to aggregate all the findings from different studies. Hence, as part of the scoping study, there is no attempt made to assess the quality or weight of the respective publications in the study basis. The main narrative account of the findings from the data charting of the publications in the study basis is collated, summarized and presented in Ingvarson (2021). In this paper, it is concluded that with only 31 publications addressing the effect of standardization in risk management regulations out of the several tens of thousands of publications covering standardization and risk in general, this specific research topic is not widely covered in scientific literature. Particularly studies of empirically measured effects from standardization of key elements of risk management and governance are largely missing. The few studies of measured effects found are mainly looking at effects of using management system standards for health, safety, and environment which are not typically considered risk management regulations.

Even with an incomplete knowledge base justifying high or low level of standardization of risk management regulations being the preferred approach, the publications include arguments related to the strengths and weaknesses of standardization. The arguments found in the scoping study are further discussed in section 4.

4. Arguments related to standardization in risk management regulations

In many of the publications found in the scoping study, arguments linked to the strengths and weaknesses of standardization and the use of standards in risk management regulations are presented. By nature of scientific literature, the arguments presented in the publications are not necessarily the viewpoint of the author(s). Often, arguments are presented and discussed without value judgements or taking sides. Publications sometimes also address aspects of standardization without framing them as arguments presented by proponents or opponents of standardization.

The most common arguments related to standardization in risk management regulations presented in the publications from the scoping study are listed in Table 2. In the table, the arguments are denoted P (pro arguments) and C (con arguments). The arguments focus on either the standardization process as such or the use of the standard that represents the output of the standardization process. This is also presented in the table.

The arguments in Table 2 are formulated to capture the essence of the argument as presented in the publications; they are not verbatim quotes from the original source. Only arguments related to standardization in risk management regulations found in more than one publication is included in the list.

5. Argument analysis

Each of the common arguments related to standardization of risk management regulations from the scoping study was subject to analysis of the arguments' strength using the approach described in section 2.2. How the argument analysis was performed in practice is further detailed in section 5.1 while section 5.2 describes how potentially contradicting arguments are managed in the argument analysis. The results from the argument analysis are presented in section 5.2.

5.1. Performing the argument analysis

Considering the inherent subjectivity of argument analysis, assessing argument strength may be significantly influenced by the assessor's knowledge and interpretation of the context in which the argument is situated. In this argument analysis, assessments were performed by two assessors with a combined 30 + years' experience as risk researcher and practitioner in the area or risk management regulations related to safety management and land-use planning. The assessors' are experienced users of risk management standards and currently involved in a research study of standardization processes through participation in standardization working groups and technical committees.

To assess the arguments' strength, each argument was analysed in terms of soundness and relevance using the assessment criteria described in section 2.2. To illustrate how each argument was analysed, consider the argument *"Best available knowledge"* (P10 in Table 2). In this argument, expert knowledge being stored in rules and technical solutions, i.e. the standard, can be considered the premise that leads to the standard representing best available knowledge. Hence, the desired quality 1 (a conclusion inevitably follows from its premises) for the *soundness* dimension is assessed as "Clearly fulfilled". Considering that standardization processes include experts in the field, there are strong reasons for believing the premise is true, so desired quality 2 is assessed as "Clearly fulfilled". Since empirical data backing this argument is available and adequately described in at least one the publications that discusses the argument, desired quality 3 is also assessed as "Clearly fulfilled". For this specific argument, all desired qualities are marked "Clearly fulfilled" resulting in labelling argument P10 *Strong* on *soundness*.

Assessing the *relevance* dimension for the same argument includes judging if premises are contextually connected with the subject or point at issue (desired quality A). For this argument, this is the case since there is a clear connection between experts defining the standard and the standard being good. Hence, desired quality A is assessed "Clearly fulfilled". There is also a pertinent substantial relation between expert knowledge stored in the standard (the premise) and the conclusion of the standard representing best available knowledge (the solution). This leads to desired quality B being assessed as "Clearly fulfilled". The evidence presented in one of the publications justifies this claim meaning that also desired quality C is assessed as "Clearly fulfilled". Argument P10 is therefore labelled *Strong* in terms of *relevance*.

To further clarify how the analysis was performed, one of the arguments rejecting standardization is used for illustration. It is argued that

standardization suffers from challenges in terms of legitimacy due to *"Inadequate transparency"* (argument C8 in Table 2). In this argument, inadequate transparency is the premise leading to legitimacy challenges (the conclusion). Applying the same criteria, the argument is assessed as *Moderate* for the *soundness* dimension. This is due to that legitimacy challenges can be seen as a result from inadequate transparency, leading to desired quality A being assessed as "Clearly fulfilled"). There are strong reasons for believing that standards suffer from inadequate transparency since many standardization processes, including the discussions and working material, is only open to the participants and not the public. Hence, desired quality B is assessed as "Clearly fulfilled". In terms of available relevant empirical source material/data, some data and anecdotal evidence can be found in the publications discussing this argument. With no empirical evidence supporting the claims being made and the non-empirical source material not being of high quality, desired quality C is assessed as "Neither clearly fulfilled nor clearly not fulfilled".

Assessing the relevance dimension for argument C8 results in desired quality A being assessed as "Clearly fulfilled" since the premise of inadequate transparency is connected to the issue of legitimacy. Since it can be argued that inadequate legitimacy can be caused by a non-transparent standardization process, desired quality B is also assessed as "Clearly fulfilled". In the publications discussing the issue of legitimacy, arguments are mainly based on intellectual reasoning rather than concrete or empirical evidence. This means that it cannot be concluded that the evidence presented have bearing on the truth of the claim at issue. Neither can it be concluded that evidence presented does not have bearing on the claim so desired quality C is assessed as "Neither clearly fulfilled nor clearly not fulfilled", resulting in a *Moderate* label for argument C8 on the *relevance* dimension.

5.2. Contradicting arguments

In the argument analysis, the desirable qualities 1–3 and A-B (see section 2.2 for details) were first assessed for each argument individually regardless of any potentially contradicting arguments or disputing evidence published in other sources. However, when combining the assessments into a Strong/Moderate/Weak label for each of the soundness and relevance dimensions (using the criteria specified in Table 1), contradicting arguments and opposing evidence should be considered since they may influence the assessment of argument strength. For example, an argument first assessed as strong based on how well it is presented in the supporting publications may be met with an equally strong counterargument found in other publications. Logically, an argument would not be assessed as strong if challenged by an even stronger counterargument or if opposing evidence is presented.

Contradicting arguments are in this argument analysis treated on a case-by-case basis, qualitatively assessing the strength of both the argument and counterargument in order to find a resulting "final" assessment of argument strength. Argument P5 *Provide cost-effective solutions* is directly contradicted by argument C4 *Inflexible and costly solutions*. Both arguments suffer from lack of empirical evidence supporting the respective claims and are primarily based on intellectual reasoning, resulting in a Weak label for the soundness dimension in the overall assessment for both arguments. Both arguments are individually labelled Moderate on the relevance dimension, an assessment maintained in the overall assessment since contradicting arguments does not impinge on how the premises of the arguments are linked to the conclusions and how they are contextually connected with the point at issue. The same reasoning is valid for contradicting arguments P6 *Regularly updated* and C3 *Fails to keep up with changes*, resulting in an overall assessment of Weak on soundness and Moderate on relevance. On the topic of standardization's impact on bureaucracy (argument P8 *Decreased administrative burden* versus C6 *Increased bureaucracy*), it is concluded that the substantial empirical evidence from several studies (e.g. Aas and Johnsen (2007) and Antonsen et al (2012)) showing "over-administration" as a result of standardization out-weigh the high-level and non-tangible reasoning behind alleged decrease in administration. Therefore, P8 is downgraded to Weak on both soundness and relevance while C6 maintains a Strong label on both dimensions.

A particular case found in the argument analysis is related to argument P10 *Best available knowledge*. This argument favouring standardization is based on the premise that a standard developed by a group of subject matter experts, the standard represents the best available knowledge. In the publications referring to this argument, this is presented as an intuitive conclusion based on intellectual reasoning without any supporting empirical evidence. Other publications, e.g. Kica and Bowman (2012), Aven and Ylönen (2019), and Lindøe and Baram (2020), disagree with this premise, claiming that standards are often developed without the representation of relevant stakeholders such as risk scientists, worker representatives and other experts. One effect of inadequate stakeholder representation (also treated separately as argument C in Table 2), it is argued, is that the standards do not represent best available knowledge, due to risk scientists and other experts not being represented in the standardization process. Therefore, P10 is down-graded to Weak on both soundness and relevance since the premise ("all" expertise being available) is not necessarily true and it cannot be concluded that the standard then represents best available knowledge.

During the analysis, a number of seemingly contradicting arguments were identified. Following careful review, however, it was concluded that they are related (sometimes even interrelated) but not in fact contradicting, resulting in an assessment of argument strength on their own merit, i.e. not considering the strength of a counterargument. This applies to arguments P1 *Contribute to uniformity* and C1 *Produce conformity and homogeneity* where P1 focus on the advantage on using a similar approach and format for managing risk in well-defined, non-complex technical or sociotechnical systems while C1 addresses challenges with standardization hampering creativity and adaption to specific needs for managing complex risks. The same distinction of how standardization comes with different effects if applied to managing conventional well-known risks and complex and emerging risks is made between arguments P2 *Increased likelihood of identifying relevant risks* and C10 *Does not address complex and emerging risks*, leading to the two arguments being assessed separately. Hence, it is important that these arguments are treated as being applicable only within each particular context as stated above.

5.3. Results from argument analysis

The arguments analysis shows (see Table 2) that arguments commonly found in the scoping study are related to both the standardization process (12 arguments) and the use of standards (14 arguments). There is no clear tendency of pro or con arguments being related to the standardization process or the use of standards.

The strength of each argument in terms of soundness and relevance is detailed in Appendix B. An overview of the argument analysis results in the appendix is presented in Fig. 2 and Fig. 3. The figures indicate that the arguments in general are assessed as stronger in terms of relevance than soundness. 21 arguments are labelled Strong or Moderate on *relevance* compared to 16 with the same label on *soundness*. This means that most arguments used in the discussion on standardization of risk management regulations being beneficial may be viewed as contextually relevant and properly relates to pros and cons of standardization.

With ten arguments labelled Weak on soundness, arguments score worse on this dimension compared to how arguments score on relevance (only five labelled Weak). Less sound arguments tend to decrease the acceptability of a standpoint since they fail to build trust in the support of a claim. Arguments being weak in terms of soundness applies to both pro and con arguments.

The type of arguments regarded as the strongest, i.e. the arguments labelled strong on both soundness and relevance, are mainly arguments opposing standardization of risk management regulations. The arguments labelled weak on both dimensions of argument strength are both pro and con arguments.

Soundness	Strong		P13	P1, P12 C1, C2, C6*, C7, C12, C13
	Moderate		P2, P9 C8, C9, C11	P3, P11
	Weak	P7, P8*, P10* C5, C10	P5*, P6* C3*, C4*	P4
		Weak	**Moderate**	**Strong**
		Relevance		

Fig. 2. Matrix with argument analysis results. Contradicting arguments are marked with * (P5/C4, P6/C3, P8/C6, P10).

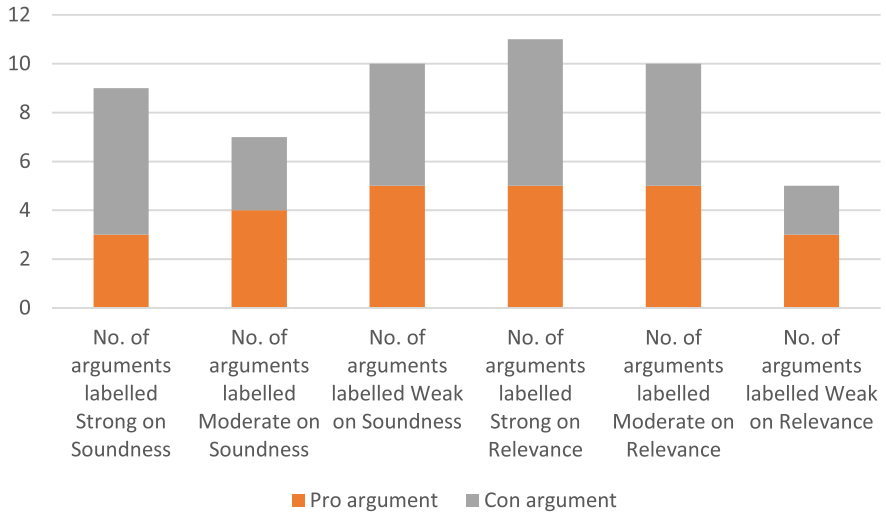

Fig. 3. Summary of argument analysis results.

Looking deeper into the argument analysis results, Fig. 4 shows that there are both similarities and differences related to the desirable qualities as illustrated in Fig. 4. Without rejecting some noticeable similarities (e.g. there are more desirable qualities assessed "Clearly fulfilled" than "Clearly not fulfilled" and "Neither clearly fulfilled nor clearly not fulfilled" combined), Fig. 4 indicates that the main desirable quality contributing to less strong arguments is *3. Argument is supported by evidence in the form of empirical source material/data OR non-empirical source material/data of high quality*. Also, desirable quality *C. Evidence presented have bearing on the truth of the claim at issue* is not fulfilled more often than the other desirable qualities. Both qualities are related to the rigour of the data supporting the claim, i.e. if the argument is adequately supported by empirical evidence or non-empirical evidence of high quality.

6. Discussion

The aim of this paper is to (1) provide an overview of the knowledge base related to effects of standardization and the use of standards in risk management regulations in the scientific literature and (2) to assess the strength of the arguments promoting or opposing standardization. The key focus of this paper is to analyse arguments presented in scientific literature related to the range of effects standardization may have rather than inserting value judgements on whether standardization is desirable. Through the systematic argument analysis, this study contributes to the debate on the most adequate approach to standardization in risk management regulations. The scoping study suggests that effects of standardization of risk management regulations are not widely covered in scientific literature. This is in line with findings from similar studies of effects of standardization in other disciplines than risk, e.g. economic impacts of standards and standardization (Blind, 2013) and the interrelationship between standards and innovation (Swann, 2000; Swann, 2010).

Potential lack of comprehensiveness, relevance, and selection bias are all inherent challenges of using literature reviews as research method (Haddaway et al., 2020). To meet these challenges, the scoping study approach was designed to include a structured search strategy and multiple databases as well as hand-searching, reference lists' review and consulting experts. The considerable share of selected publications stemming from the consultation exercise (17 of 31) may indicate an inadequate search strategy. However, a large share of the publications from the database search were also suggested by the experts, stressing the value of the consultation exercise being part of the scoping study design to ensure scientific rigour. The "accuracy" of the publication suggested by experts, i.e. the number of selected publications in relation to the number of candidates for abstract review, is significantly higher than the more mechanical data base searching. Also, eight of the 17 publications suggested by experts were book sections from the same publication, resulting in further skewed representation in the final selection of publications. Despite a limited knowledge base in academic literature, an active discussion on the value of standardizing risk is ongoing. The argument analysis captures many arguments advocating and opposing standardization of risk management regulations. The arguments are related to both the standardization process as such and the standards that represent the output of these processes, suggesting that standards and the process of standardization should be seen as two inseparable aspects of standardization of risk management regulations.

A fundamental challenge when assessing the strength of arguments, a topic of discussion among scholars since the early days of science, is the inevitably inherent subjectivity and the limited knowledge of the evaluator. To limit, but not eliminate, the issue of subjectivity, bias and arbitrariness, an argument analysis approach and argument assessment

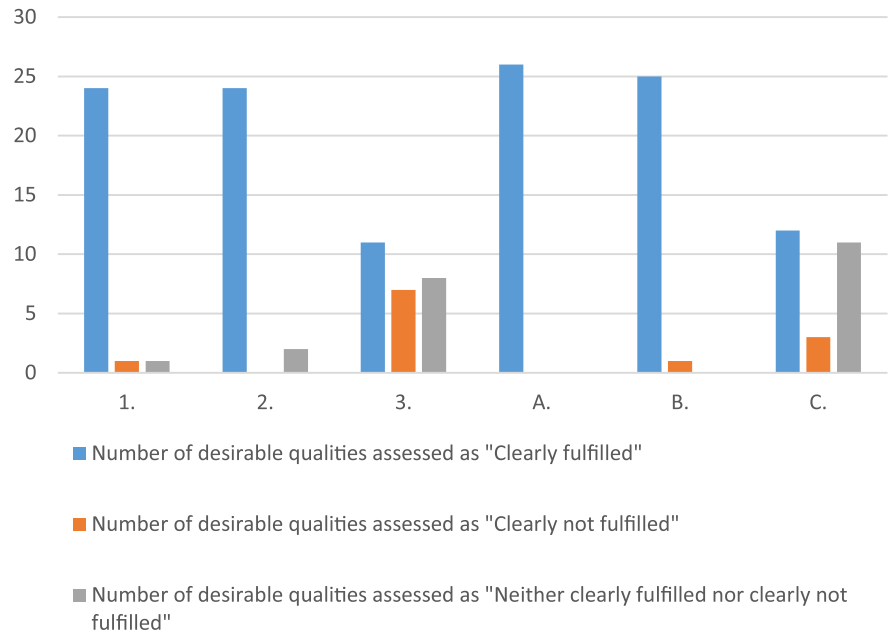

Fig. 4. Summary of argument analysis results. Qualities 1–3 are related to the Soundness dimension and A-C are related to the relevance dimension of argument strength (see section 2.2 for more details).

criteria were developed as part of this paper using inspiration from both argumentation theory and risk research. Even so, subjectivity remains and the assessment of argument strength could be challenged and debated and the strength of the arguments should continuously be reconsidered in the light of new information and knowledge.

In this paper, the arguments are often assessed as less strong in terms of soundness than relevance. This means that even if the arguments are not successful in building trust and increasing the acceptability of the claim or standpoint, the argumentation as such is contextually relevant and a good starting point for a dialectic on the topic. With arguments being relevant, addressing the strengths and weaknesses of standardization of risk management regulations will help establishing an appropriate approach to standardization in risk management regulations through reasoned argumentation.

The perceived less strong soundness dimension of the pro and con arguments can be explained by the lack of empirical evidence backing the argument. The finding of limited empirical evidence being available is not unexpected nor unusual. Safety science, which is often considered closely linked to risk science, has been criticized for favouring publications on theory and methods rather than empirical evidence production and intervention evaluations (Rae et al., 2020). To mitigate this, empirical studies related to effects of standardization and the use of standards in risk management regulations are required to substantiate argument strength and help build trust in both the arguments and the justification of the chosen risk regulatory regime.

An important finding of the argument analysis performed in this study is that some arguments are related to the same issue but presented in opposing ways, i.e. the same argument is used both as pro and con. For example, standardization decreasing the administrative burden (argument P8 in Table 2) is contradictory to the argument that standardization increases bureaucracy (C6). It can be argued, as done in section 5.2, that an argument will lose its strength if the same argument is used as counter argument. However, it can also be argued that argumentation and arguments are so contextual that the same argument can be seen as counter arguments with its own merits in terms of strength and thus not affected by self-contradiction. This means that standardization may be fairly seen as increasing the administrative burden if the standard impose more elaborate and time-consuming administrative requirements than what would be the case without the standard. On the other hand, due to the predictable nature of a standard, an adequately designed standard may also provide stream-lining opportunities for previously "over-administrated" organizations or procedures. This calls for additional research aiming at more context-specific evaluations where the merits of standardization is analysed with specific attention to how standardization processes are implemented, with what objectives, at what level of detail, and with what effects seen from the perspective of different stakeholders.

Another aspect that potentially affects argument strength is if the same or similar argument is presented in different publications using different underlying information. Intuitively, arguments from different sources based on various data will increase the argument's strength. However, here each argument is individually assessed in terms of soundness and relevance, so "scores" are not accumulated. This means that an argument does not benefit from being mentioned by many. Even if the same argument is found in different publications, it will still be assessed as a weak argument if the basis supporting the argument is unsound or irrelevant. It is the particular assessment of soundness/relevance in each context that underpins argument strength. If the same argument is assessed differently in different publications, the highest assessment is used to represent that argument's strength in Fig. 2. For example, an argument found in three different publications would be labelled Moderate in case the argument is assessed Moderate once and Weak twice in the different publications. Consequently, an argument assessed Weak in several different publications would be labelled Weak.

An alternative approach for argument analysis would be to define one argument and analyse how this argument is underpinned in different publications. Intuitively, if different aspects supporting the argument is presented in several publications, the argument would be considered stronger than if only one and the same aspect was repeated in all publications. This kind of approach could also shed more light on how the same argument is built in different contexts, potentially strengthening the argument. This type of argument analysis would require a different synthesis than what is used in this paper where arguments have been singled out from publications selected from a literature review. Hence, this alternative approach may be a valuable option for future research efforts.

In the chosen approach for argument analysis, empirical evidence is a prerequisite for sound and strong arguments. It may be argued that for some type of arguments, the logical inference and intellectual reasoning is so solid that empirical evidence may not be explicitly required to

constitute a sound argument. For example, standards being a way to demonstrate compliance (argument P11 in Table 2) and contributing to uniformity (P1) are inherent features of standardization in its semantic meaning of evaluating something by comparing it to an established norm or standard (Juhl, 2020). Even if strict application of the criteria for classifying argument strength may indicate that this type of arguments is assessed as less sound or strong, arguments may still be valid even if not backed by explicit empirical evidence.

Another key aspect of argument strength is the premises and their relation to the conclusion of the argument (i.e. relevance). Some premises may seem to have an intuitively strong relation to the conclusion but when scrutinized, the premise can be challenged. Take for example argument P10 in which standards represent "best available knowledge". The premise here is that experts are gathered to design the standard leading to the conclusion that their collective expertise is captured in the standard. But is it reasonable to assume that "all" experts are gathered? And that their respective expertise is collected without conflicts and negotiations resulting in solutions that the participants can live with rather than representing what they think is the best solution?

The argument analysis shows that many of the arguments are assessed as rather weak. This may indicate a general need for additional studies that can be used to provide support for or against the arguments. In addition, it may indicate a need for more in-depth reviews of the existing knowledge base. Furthermore, although many arguments are assessed as weak, the arguments as such should not necessarily be disregarded as being objectively poor. Some arguments may be valid even if empirical data or similar supporting evidence is not available in the specific publication included in the scoping study. For example, argument P9 that standardization may improve the use of resources and expertise by sharing expert knowledge across organizations is both reasonable and intellectually intuitive even if no empirical or non-empirical data supporting this statement is presented. Similarly, standardization processes may suffer from inadequate stakeholder representation (argument C9) even if relevant empirical source material/data were not available in the specific publications discussing this topic. Therefore, the arguments assessed as weak in this paper should not be disqualified. Instead, all arguments should be individually assessed and discussed on case-by-case basis when using the argument in specific contexts with specific objectives.

Many of the arguments are interconnected and affect each other to some extent. It seems natural that standards representing best available knowledge (P10) and are regularly updated (P6) provide cost-effective solutions (P5) and increase the safety of products and safety for humans and the environment (P4). But if standards fail to keep up with changes and becomes outdated (C3), they may be less flexible leading to less effective risk-reducing measures and costly solutions (C4) and instead act as a barrier to innovation (C7). And if the standard does not encourage higher ambition levels than the specified minimum solution (C2), this may further decrease the motivation to perform risk analyses (C13) and spur deterioration of technical competence and local knowledge (C12). This interconnectedness between arguments and counter-arguments demonstrates the challenge of arriving at a conclusion if standardization of risk management regulations is mainly beneficial or not. To further advance the field through a more detailed study of argument interconnectedness, follow-up research may include the use of correlation matrices, network analysis or other tools to depict and analyse relations between arguments. This would potentially shed more light on how pro and con arguments relates to each other.

Linked to the interconnectedness between arguments is the "chicken or the egg" challenge of one argument resulting in another argument, e.g. is the challenge of standards becoming outdated due to new knowledge (C3) a result of inadequate stakeholder representation (C9) or a stand-alone argument? This challenge seems less significant as long as all relevant arguments found in the publications from the scoping study are captured in one argument or another in the complete list of arguments (Table 2). The argument strength is likely to be assessed the same since the basis for the argument remains the same no matter how the argument is categorized. Since argument strength is analysed independently of any potential counterarguments, a valuable next level of analysis could be the more holistic synthesis of arguments discussed above rather than the current assessment based on a "paper by paper" approach.

Discussing the value of standardization includes a strong link to achieving objectives. In case the overarching objective is to achieve a uniform and potentially resource efficient risk assessment approach, standardization may be seen as beneficial. If the aim is to provide a comprehensive and accurate characterization of risk, assigning the level of risk a numerical value and evaluate its tolerability, standardization may be less beneficial due to the standard's shortcomings in terms of not representing a common set of norms for all stakeholders. A standardized approach then positively impacts achieving the former objective but negatively impacts achieving the latter objective. This suggests that preferring or rejecting standardization is sometimes more a question of which objectives to achieve in a given context than on ideological preferences or diverse opinions on argument strength in general.

As discussed above, benefits and disadvantages of standardization of risk management regulations are highly contextual. The benefits of standardization are directly linked to the situation in which the standard is used and how the standard is designed. A mass-produced conventional, non-complex technical system is likely to benefit from standardization since it brings predictability and resource-efficiency in design, implementation and operation. At the same time, emerging risks (e.g. IRGC, 2017) such as complex socio-technical systems including novel technology and challenging interfaces may be hindered by standards that are developed based on past experiences and situational awareness that are no longer relevant. This is one reason why two seemingly opposing arguments can be perceived as strong; they may both be valid but refer to different contexts where the standardisation process may have been implemented in different ways.

A perspective of standardization in risk management regulations not prominent in the present argument analysis is that of politics and power. When using the selection criterion of an argument requiring to be found in more than one publication, no political arguments were selected for argument analysis, even if the political perspective is highlighted in some publications (e.g. by Lindøe and Baram (2020) and Thompson (2020)). This is as natural reflection of the search string used for the database search which was developed with a specific focus of effects of standardization rather than what influences standardization in terms of balance of interest and power. Political considerations are an integral part of how risk management regulations are designed, i.e. level or degree of standardization, and intrinsic in the development of standards. Political positioning may also impact the assessment of argument strength. Addressing the perspective of politics and power is important but as a result of this study's limitations and methodological approach, this is not covered in this paper. Instead, this perspective is the dedicated focus of an ongoing research activity aimed at exploring the politics of standardization processes and the power relations exercised in standardization bodies.

The argument analysis focuses on the benefits and challenges with standardization mainly in terms of degrees of standardization of performing risk assessments, i.e. prescriptive versus non-prescriptive risk assessment procedures. This means that mainly a subset (the risk assessment) of the risk management process is scrutinized. Considering the distinction between *process* and *outcome* standards (Brunsson et al., 2012), standardizing the overall risk management and risk governance processes has not been the focus of this study but may be worth exploring in terms of strengths and weaknesses. Also, the required level of guidance provided in the standard should be examined in future research since many standards, particularly some key ISO standards (Björnsdóttir et al., 2021), come with inadequate guidance for implementation. This would provide additional valuable input to the knowledge base when assessing the most adequate approach to

standardization in risk management regulations.

The limited number of publications with argumentation based on empirical studies of standardization of risk management regulations may be due to difficulties of measuring the actual effects of standardization as an intervention. Evaluating, and even more so measuring, the effects of interventions are challenging from many perspectives (see e.g. Hassel and Cedergren, 2021). A main complication is to distinguish standardization from a multitude of other potential influencing factors as the cause of certain, potentially measured, effects. Standardizing risk management and risk governance approaches also suffer from the same challenge as other safety interventions: they have delayed effects. For example, deterioration of technical competence will not occur instantly. Neither will the effects of standardization of risk management regulations by helping facilitate the integration of several sources of information in multi-actor risk assessment processes be observed over-night.

To meet the challenges of evaluating the effects of standardization of risk management regulations, the effects of such interventions must be isolated. This can be achieved using a structure that clarifies which mechanisms lead to certain effects or outputs by the use of a specific intervention in given circumstances. One option that may be beneficial to pursue as part of future research efforts is to use CIMO-logic (Denyer et al., 2008). CIMO-logic is a design science research tool aiming at designing interventions to solve problems. If used reversely, the CIMO-logic may provide a systematic structure to evaluate how certain Interventions (I) using determined Mechanisms (M) deliver specific Outcomes (O) in certain Contexts (C).

7. Conclusions

The main conclusions from this scoping study and argument analysis are that effects of standardization of risk management regulations are not widely covered in scientific literature and empirical evidence is largely lacking to support arguments advocating and opposing standardization of risk management regulations. Although not often underpinned by empirical evidence, there are, however, strong opinions and multiple arguments promoting and opposing standardization. Assessing if the strengths of standardization outweigh the weaknesses is subjective and highly contextual. Even if the argument analysis presented in this paper indicates that many of the arguments are labelled weak, they should not be disregarded since the arguments may be logically correct and intuitively reasonable even if they lack appropriate empirical backing in the particular publication(s).

As discussed in this paper, argument analysis includes significant elements of subjectivity that cannot be eliminated entirely by the use of a rigorous approach. This may lead to potential disagreement on details of the analysis results in terms of individual arguments being correctly being labelled strong, moderate or weak. Considering that the strength of the arguments should continuously be reconsider in the light of new information and knowledge, this argument analysis should be viewed not as an end point but a starting point for discussing argument strength. And with some of the arguments also being interconnected and contradictory, it seems not possible to come to a conclusion and reach an agreement in the discussion on standardization of risk management regulations being beneficial. Thus, the discussion should be in the form of dialectic, i.e. critically reflect and reason by means of dialogue, rather than debate with its negative connotations of quarrel and disagreement.

Acknowledging, understanding and addressing the various arguments in favour of and rejecting standardization, as well as their respective context, is essential when designing risk management regulations. To develop a more comprehensive knowledge base, additional empirical research is required on the effects of standardization of risk management regulations. CIMO-logic and similar approaches may prove valuable for this by offering a structure for isolating the intervention from other influencing factors. Future research efforts should also be directed towards exploring *when* (i.e. in which situations) and *how* the risk management and risk governance processes or their components can be standardized and the appropriate *level* of standardization in terms of prescriptive regulations. Also, the political and power perspectives of standardization of risk management regulations and the politicized standardization process and standardization organizations may be valuable to explore in future research.

CRediT authorship contribution statement

Johan Ingvarson: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Henrik Hassel:** Writing – review & editing, Supervision, Methodology, Formal analysis, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research has been funded by the Swedish Transport Administration. The funding source has had no role in study design, in the collection, analysis and interpretation of data nor in the writing of the paper. The authors are grateful to three anonymous reviewers for useful and constructive comments on earlier versions of the manuscript.

This research has been funded by the Swedish Transport Administration. The funding source has had no role in study design; in the collection, analysis and interpretation of data nor in the writing of the paper.

Appendix A. Inclusion and exclusion criteria

The set of inclusion and exclusion criteria used in the database search is provided in Table 3.

Table 3
Inclusion and exclusion criteria.

ID	Inclusion (I) / Exclusion (E) criteria
I1	Full text publication available in English.
I2	Publication is peer-reviewed.
I3	Publication date is from year 2000 to date.
I4	Publication is related to (and in the database labelled with) a relevant subject area or research category.
I5	Publication should include all key components of the research question.
I6	Publications should discuss the effects or experiences of using standards or standardized approaches; not the use of standards or standardized approaches itself.
E1	Answer is no to I1, I2 or I3.
E2	Publication is exclusively related to (and in the database labelled with) an irrelevant subject area or research category.
E3	Publication discusses the implementation of a single preferred standard or standardized approach in a specific situation.
E4	Publication discusses standard operating procedures or standardized approaches used in medicine or similar disciplines.

Appendix B. Strength of argument assessment

The assessments of the strength of each argument in terms of soundness and relevance is are provided in Table 4. The soundness dimension includes the following desirable qualities:

1. A conclusion inevitably follows from its premises.
2. There are strong reasons for believing the premises are true.
3. Argument is supported by evidence in the form of empirical source material/data OR non-empirical source material/data of high quality.

Table 4
Assessment of argument strength in terms om soundness and relevance.

ID	Argument	Brief description	Argument focus	Source publication	Soundness				Relevance			
					Desirable qualities			Label	Desirable qualities			Label
					1.	2.	3.		1.	2.	3.	
P1	Contribute to uniformity	Contribute to uniformity (e.g. defining, identifying, analysing, assessing and communicating risk) making risk more comparable and therefore more manageable.	Standard	Aven and Ylönen (2019) Olsen (2020b) Kica and Bowman (2012) Antonsen, et al. (2012) Aas and Johnsen (2007) Jore and Moen (2015)	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong
P2	Increased likelihood of identifying relevant risks	Support the identification and inclusion of all relevant risks, including those not of particular interest to the party performing the analysis.	Standard	Aven and Ylönen (2019) Aas and Johnsen (2007) do Nascimento, et al. (2020)	Neither clearly fulfilled nor clearly not fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Clearly fulfilled	Moderate	Clearly fulfilled	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate
P3	Bench-marking opportunities	Compare practices with internationally recognised benchmarks and allows a common approach to be taken across boundaries (i.e. a single international reference point).	Standard	Aven and Ylönen (2019) MacRae (2011)	Clearly fulfilled	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong
P4	Increased safety	Increase the safety of products and safety for humans and the environment.	Standard	Aven and Ylönen (2019) do Nascimento, et al. (2020) Murakami (2016) MacRae (2011)	Clearly fulfilled	Clearly fulfilled	Clearly not fulfilled	Weak	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong
P5	Provide cost-effective solutions	Using standards, there is a potential for lower costs when there is no need to demonstrate compliance or explore alternative solutions.	Standard	Nyvik, et al. (2021) Belluck, et al. (2006)	Clearly fulfilled	Clearly fulfilled	Clearly not fulfilled	Weak	Clearly fulfilled	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate
P6	Regularly updated	Standards are regularly reviewed and updated providing opportunities to introduce recent developments.	Standard	Aven and Ylönen (2019) Kica and Bowman (2012)	Clearly fulfilled	Clearly fulfilled	Clearly not fulfilled	Weak	Clearly fulfilled	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate
P7	Offers predictable low risk levels	Provide lower risk levels by offering predictability (i.e. if a standard is complied with, a predicted adequate risk level is achieved).	Standard	Thompson (2020) do Nascimento, et al. (2020)	Clearly not fulfilled	Clearly fulfilled	Clearly not fulfilled	Weak	Clearly fulfilled	Clearly not fulfilled	Clearly not fulfilled	Weak
P8	Decreased administrative burden	Provide predictable and predefined targets, approaches and procedures to which the organization can streamline internal work processes.	Standard	Kica and Bowman (2012) Belluck, et al. (2006) MacRae (2011))	Clearly fulfilled	Clearly fulfilled	Clearly not fulfilled	Weak	Clearly fulfilled	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate
P9	Effective allocation and use of resources and expertise	Prevents “re-inventing the wheel” allowing organizations to capitalize on the expertise/resources of others rather than requesting all organizations to have all expertise/resources available in-house.	Standardization	Aven and Ylönen (2019) Foliente (2000) Lindøe and Baram (2020)	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate	Clearly fulfilled	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate
P10	Best available knowledge	Expert knowledge is stored in rules and technical solutions representing “best available knowledge”.	Standardization	Nyvik, et al (2021) Skotnes, 2020Lindøe and Baram (2020) Antonsen, et al (2012)	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong
P11	Compliance verification	Standards are related to expectations of governance and represents defendable,	Standardization	Skotnes, 2020Foliente (2000) Aven and Ylönen	Clearly fulfilled	Clearly fulfilled	Neither clearly	Moderate	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong

(*continued on next page*)

Table 4 (*continued*)

ID	Argument	Brief description	Argument focus	Source publication	Soundness				Relevance			
					Desirable qualities			Label	Desirable qualities			Label
					1.	2.	3.		1.	2.	3.	
		auditable processes and a way to demonstrate compliance. Standards are easy for third party to check and regulators to enforce.		(2019) Belluck, et al. (2006) Jore and Moen (2015)			fulfilled nor clearly not fulfilled					
P12	Facilitate risk governance	Help facilitate the integration of several sources of information in multi-actor risk assessment processes and reduce flexibility in risk governance.	Standardization	Tehler, et al (2020) Olsen (2020a) Aas and Johnsen (2007)	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong
P13	Hegemonic approach to managing risk	Contribute to building a legitimate hegemonic approach to the management of risk that is especially helpful for topic areas that are less often recognized and where organizations have little experience.	Standardization	Jore, 2020Aas and Johnsen (2007)	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong	Clearly fulfilled	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate
C1	Produce conformity and homogeneity	Standardization produces conformity and homogeneity between stakeholders, while risk management and governance requires diversity and heterogeneity.	Standard	Aven and Ylönen (2019) Olsen (2020a) Jore, 2020Jørgensen and Lindøe (2020)	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong
C2	Compliance-driven	Specifies minimum solutions and does not encourage higher ambition levels. Compliance with the standard(s) is in focus, not the optimal level of risk.	Standard	Nyvik, et al (2021) Clark-Ginsberg and Slayton (2019) Jore and Moen (2015)	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong
C3	Fails to keep up with changes	The standard fails to keep up with changes and becomes outdated due to new knowledge or new alternative solutions becoming available/ developed.	Standard	Nyvik, et al (2021) Clark-Ginsberg and Slayton (2019) Aas and Johnsen (2007) Jore and Moen (2015)	Clearly fulfilled	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate	Clearly fulfilled	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate
C4	Inflexible and costly solutions	The standard may not be developed for the context specific conditions and appear less flexible which may result in less effective risk-reducing measures and costly solutions.	Standard	Nyvik, et al (2021) Raj and Lemoff (2009) Foliente (2000) Jore and Moen (2015)	Clearly fulfilled	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate	Clearly fulfilled	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate
C5	Universally agreed norms not available	Normative judgements are at the core of both risk and standardisation and since there are no/few universally agreed set of norms, standardisation is challenging. This also includes the discussion on what level of risk is acceptable and to whom?	Standard	Juhl, 2020MacRae (2011)	Clearly fulfilled	Clearly fulfilled	Clearly not fulfilled	Weak	Clearly fulfilled	Clearly fulfilled	Clearly not fulfilled	Weak
C6	Increased bureaucracy	Standards may include administrative requirements that are cumbersome to some organizations (typically smaller organizations with less comprehensive management systems).	Standard	Antonsen, et al. (2012) MacRae (2011) Aas and Johnsen (2007)	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong
C7	Barrier to innovation	Barrier to innovation by locking the industry into compliance with 'one-size-fits-all' rules, thereby disincentivizing and limiting flexibility and adaption to new technologies.	Standardization	Aven and Ylönen (2019) Foliente (2000) Lindøe and Baram (2020) Onofrio, et al. (2015)	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong
C8	Inadequate transparency	Challenges in terms of legitimacy due to inadequate transparency (i.e. public	Standardization	Kica and Bowman (2012) Lindøe and Baram (2020)	Clearly fulfilled	Clearly fulfilled	Neither clearly	Moderate	Clearly fulfilled	Clearly fulfilled	Neither clearly	Moderate

(*continued on next page*)

Table 4 (*continued*)

ID	Argument	Brief description	Argument focus	Source publication	Soundness				Relevance			
					Desirable qualities			Label	Desirable qualities			Label
					1.	2.	3.		1.	2.	3.	
		access to information and decision-making procedures).					fulfilled nor clearly not fulfilled				fulfilled nor clearly not fulfilled	
C9	Inadequate stakeholder representation	Challenges in terms of legitimacy due to inadequate or imbalanced stakeholder representation (e.g. involvement of experts and presence and participation of unions, workers, the public, other industries, etc.).	Standardization	Aven and Ylönen (2019) Kica and Bowman (2012) Lindøe and Baram (2020) Urciuoli (2018)	Clearly fulfilled	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate	Clearly fulfilled	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate
C10	Does not address complex and emerging risks	Complex risks are associated with limited historical data (i.e. large uncertainties) which is a typically required for successful standardization. This may also reduce comparability with other complex risks and could reduce the capacity to discover and understand new risks on the horizon.	Standardization	Jore, 2020Olsen (2020a)	Clearly fulfilled	Clearly fulfilled	Clearly not fulfilled	Weak	Clearly fulfilled	Clearly fulfilled	Clearly not fulfilled	Weak
C11	May generate false sense of safety	Standardization may reduce our attention to unthinkable threats and risks it is easy to talk about but difficult to handle, potentially generating a false sense of safety/security.	Standardization	Skotnes, 2020Olsen (2020a) Jore, 2020Clark-Ginsberg and Slayton (2019)	Clearly fulfilled	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate	Clearly fulfilled	Clearly fulfilled	Neither clearly fulfilled nor clearly not fulfilled	Moderate
C12	Deterioration of technical competence	Reliance on industry standards and practices may lead to deterioration of own technical competence and local knowledge.	Standardization	Lindøe and Baram (2020) Antonsen, et al. (2012)	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong
C13	Reduced ownership and quality	Standardization may lead to a loss of ownership and reduced motivation to perform risk analyses (or other risk management activities) and thus reduce the quality of assessments/ activities.	Standardization	Tehler, et al. (2020) Olsen (2020a) Antonsen, et al. (2012)	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong	Clearly fulfilled	Clearly fulfilled	Clearly fulfilled	Strong

The relevance dimension includes the following desirable qualities:

A. Premises are contextually connected with the subject or point at issue.
B. There is a pertinent substantial relation between the premises and the conclusion of the argument.
C. Evidence presented have bearing on the truth of the claim at issue.

References

Aas, A.L., Johnsen, S.O. (2007). Improvement of human factors in control centre design - Experiences using ISO 11064 in the Norwegian Petroleum Industry and suggestions for improvements. Paper presented at the International Petroleum Technology Conference.

Antonsen, S., Skarholt, K., Ringstad, A.J., 2012. The role of standardization in safety management – A case study of a major oil & gas company. Saf. Sci. 50 (10), 2001–2009. https://doi.org/10.1016/j.ssci.2011.11.001.

Arksey, H., O'Malley, L., 2005. Scoping studies: towards a methodological framework. Int. J. Soc. Res. Methodol. 8 (1), 19–32. https://doi.org/10.1080/1364557032000119616.

Askeland, T., Flage, R., Aven, T., 2017. Moving beyond probabilities – Strength of knowledge characterisations applied to security. Reliab. Eng. Syst. Saf. 159, 196–205. https://doi.org/10.1016/j.ress.2016.10.035.

Aven, T., Ylönen, M., 2019. The strong power of standards in the safety and risk fields: A threat to proper developments of these fields? Reliab. Eng. Syst. Saf. 189, 279–286. https://doi.org/10.1016/j.ress.2019.04.035.

Belluck, D.A., Hull, R.N., Benjamin, S.L., Alcorn, J., Linkov, I., 2006. Are standard risk acceptability criteria applicable to critical infrastructure based on environmental security needs?'. In: Morel, B., Linkov, I. (Eds.), Environmental Security and Environmental Management: The Role of Risk Assessment, Vol. 5. Springer, Dordrecht, p. pp. 93-+)..

Björnsdóttir, S.H., Jensson, P., Boer, R.J., Thorsteinsson, S.E., 2021. The Importance of Risk Management: What is Missing in ISO Standards? Risk Analysis: Int. J. 1 https://doi.org/10.1111/risa.13803.

Blind, K. (2013). The Impact of Standardization and Standards on Innovation (NESTA Working Paper 13/15).

Brunsson, N., Rasche, A., Seidl, D., 2012. The Dynamics of Standardization: Three Perspectives on Standards in Organization Studies. Organization Studies 33 (5/6), 613–632. https://doi.org/10.1177/0170840612450120.

Chatfield, T., 2018. Critical Thinking: Your guide to Effective Argument, Successful Analysis & Independent Study. SAGE Publishing, London.

Clark-Ginsberg, A., Slayton, R., 2019. Regulating risks within complex sociotechnical systems: Evidence from critical infrastructure cybersecurity standards. Science and Public Policy 46 (3), 339–346. https://doi.org/10.1093/scipol/scy061.

Colquhoun, H.L., Levac, D., O'Brien, K.K., Straus, S., Tricco, A.C., Perrier, L., Moher, D., 2014. Scoping reviews: time for clarity in definition, methods, and reporting. J. Clin. Epidemiol. 67, 1291–1294. https://doi.org/10.1016/j.jclinepi.2014.03.013.

Cooper, H.M. (1985). A Taxonomy of Literature Reviews. Paper presented at the Annual Meeting of the American Educational Research Association Chicago.

Daudt, H.M.L., van Mossel, C., Scott, S.J., 2013. Enhancing the scoping study methodology: a large, inter-professional team's experience with Arksey and O'Malley's framework. BMC Med. Res. Method. 13 (1), 1–9. https://doi.org/10.1186/1471-2288-13-48.

Dent, C.J., Bell, K.R.W., Richards, A.W., Zachary, S., Eager, D., Harrison, G.P., Bialek, J. W. (2010). The role of risk modelling in the great Britain transmission planning and operational standards. Paper presented at the 2010 IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems, PMAPS 2010, Singapore.

Denyer, D., Tranfield, D., van Aken, J.E., 2008. Developing Design Propositions through Research Synthesis. Organization Studies 29, 393–413.

do Nascimento, A. P., dos Santos, W. R., & de Oliveira, M. P. V. (2020). The risk mentality in organizations: An analysis of inserting risk management in ISO 9001 and ISO 14001: 2015 standards. Gestao e Producao, 27(2). doi: https://doi.org/10.1590/0104-530X4043-20.

Ferrari, R., 2015. Writing narrative style literature reviews. Medical Writing 24 (4), 230–235. https://doi.org/10.1179/2047480615Z.000000000329.

Flage, R., Aven, T., 2009. Expressing and communicating uncertainty in relation to quantitative risk analysis (QRA). Risk Reliab: Theory Appl 2, 9–18.

Flage, R., Aven, T., 2017. Comments to the article by Goerlandt & Reniers titled "On the assessment of uncertainty in risk diagrams" [Safety Sci. 84 (2016) 67–77]. Saf. Sci. 98, 9–11. https://doi.org/10.1016/j.ssci.2017.04.007.

Flage, R., Aven, T., 2018. Some brief concluding remarks in relation to the discussion with Floris Goerlandt and Genserik Reniers about strength of knowledge (strength of evidence) judgments in semi-quantitative risk analysis. Saf. Sci. 108, 237. https://doi.org/10.1016/j.ssci.2017.08.005.

Foliente, G.C., 2000. Developments in Performance-Based Building Codes and Standards. Forest Products Journal 50 (7/8), 12.

Goerlandt, F., Khakzad, N., Reniers, G., 2017. Validity and validation of safety-related quantitative risk analysis: A review. Saf. Sci. 99, 127–139. https://doi.org/10.1016/j.ssci.2016.08.023.

Goerlandt, F., Reniers, G., 2017. Evidence assessment schemes for semi-quantitative risk analyses: A response to Roger Flage and Terje Aven. Saf. Sci. 98, 12–16. https://doi.org/10.1016/j.ssci.2017.04.008.

Grant, M.J., Booth, A., 2009. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Information & Libraries Journal 26 (2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x.

Haddaway, N.R., Bethel, A., Dicks, L.V., Koricheva, J., Macura, B., Petrokofsky, G., Pullin, A.S., Savilaakso, S., Stewart, G.B., 2020. Eight problems with literature reviews and how to fix them. Nat. Ecol. Evol. 4 (12), 1582–1589. https://doi.org/10.1038/s41559-020-01295-x.

Hassel, H., Cedergren, A., 2021. A framework for evaluating societal safety interventions. Saf. Sci. 142, 105393 https://doi.org/10.1016/j.ssci.2021.105393.

Hoffmann, M.H.G., 2018. The Elusive Notion of "Argument Quality". Argumentation 32 (2), 213–240. https://doi.org/10.1007/s10503-017-9442-x.

Ingvarson, J. (2020). The Basis for Debating Standardization of Risk for Land-use Planning is Incomplete. Paper presented at the Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference.

Ingvarson, J. (2021, 19-23 September). Standardization in Risk Management Regulations: What Can We Learn From Scientific Literature? Paper presented at the 31st European Safety and Reliability Conference (ESREL 2021), Angers, France.

IRGC. (2017). Introduction to the IRGC Risk Governance Framework, revised version.

Johnson, R.H., Blair, J.A. (2006 [1977]). Logical self-defense. New York: International Debate Education Association.

Jore, S.H., 2020. Standardization of terrorism risk analysis - A means or an obstacle to achieving security? In: Olsen, O.E., Juhl, K., Lindøe, P.H., Engen, O.A. (Eds.), Standardization and Risk Governance - A Multi-Disciplinary Approach. Routledge, pp. 150–165.

Jore, S.H., Moen, A., 2015. A discussion of the risk-management and the rule-compliance regulation regimes in a security context. In: Nowakowski, T., Mlynczak, M., JodejkoPietruczuk, A., WerbinskaWojciechowska, S. (Eds.), Safety and Reliability: Methodology and Applications. Crc Press-Taylor & Francis Group, Boca Raton, pp. 677–684.

Juhl, K., 2020. Standardization of risk versus the risk of standardization - A conceptual analysis. In: Olsen, O.E., Juhl, K., Lindøe, P.H., Engen, O.A. (Eds.), Standardization and Risk Governance - A Multi-Disciplinary Approach. Routledge, pp. 16–40.

Jørgensen, L., Lindøe, P.H., 2020. Standardizations and risk mapping: Strengths and weaknesses. In: Olsen, O.E., Juhl, K., Lindøe, P.H., Engen, O.A. (Eds.), Standardization and Risk Governance - A Multi-Disciplinary Approach. Routledge, pp. 181–198.

Kica, E., Bowman, D.M., 2012. Regulation by Means of Standardization: Key Legitimacy Issues of Health and Safety Nanotechnology Standards. Jurimetrics 53 (1), 11–56.

Kowalczyk, N., Truluck, C., 2013. Literature Reviews and Systematic Reviews: What Is the Difference? Radiol. Technol. 85 (2), 219–222.

Levac, D., Colquhoun, H., O'Brien, K.K., 2010. Scoping studies: advancing the methodology. Implementation Science 5, 69–77. https://doi.org/10.1186/1748-5908-5-69.

Lindøe, P.H., Baram, M.S., 2020. The role of standards in hard and soft approaches to safety regulation. In: Olsen, O.E., Juhl, K., Lindøe, P.H., Engen, O.A. (Eds.), Standardization and Risk Governance - A Multi-Disciplinary Approach. Routledge, pp. 235–254.

MacRae, D., 2011. Standards for risk assessment of standards: How the international community is starting to address the risk of the wrong standards. J. Risk Res. 14 (8), 933–942. https://doi.org/10.1080/13669877.2011.571782.

Mays, N., Roberts, E., Popay, J., 2001. Synthesising research evidence. In: Fulop, N., Allen, P., Clarke, A., Black, N. (Eds.), Studying the organisation and delivery of health services : research methods. Routledge, London.

Morsut, C. (Ed.), 2020. Towards a standardization of EU disaster risk management?. Routledge.

Munn, Z., Peters, M.D.J., Stern, C., Tufanaru, C., McArthur, A., Aromataris, E., 2018. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med. Res. Method. 18 (1), 1–7. https://doi.org/10.1186/s12874-018-0611-x.

Murakami, M., 2016. Risk Analysis as Regulatory Science: Toward The Establishment of Standards. Radiat. Prot. Dosim. 171 (1), 156–162. https://doi.org/10.1093/rpd/ncw211.

Nutley, S., Powell, M., Davies , H. (2013). What counts as good evidence? Alliance for Useful Evidence.

Nyvik, C.L.B., Flage, R., Guikema, S., 2021. On the Use of Standards and Guidelines as a Tool to Fulfil Regulatory Requirements. Risk Anal.: Off. Publ. Soc. Risk Anal. https://doi.org/10.1111/risa.13672.

Olsen, O.E., 2020. Dilemmas of standardization in risk governance. In: Olsen, O.E., Juhl, K., Lindøe, P.H., Engen, O.A. (Eds.), Standardization and Risk Governance - A Multi-Disciplinary Approach. Routledge, pp. 275–280.

Olsen, O.E., 2020. The standardization of risk governance. In: Olsen, O.E., Juhl, K., Lindøe, P.H., Engen, O.A. (Eds.), Standardization and Risk Governance - A Multi-Disciplinary Approach. Routledge, pp. 3–15.

Onofrio, R., Piccagli, F., Segato, F. (2015). Failure Mode, Effects and Criticality Analysis (FMECA) for medical devices: Does standardization foster improvements in the practice? Paper presented at the 6th International Conference on Applied Human Factors and Ergonomics, Amsterdam.

Peters, M.D.J., Godfrey, C.M., Khalil, H., McInerney, P., Parker, D., Soares, C.B., 2015. Guidance for conducting systematic scoping reviews. International Journal of Evidence-Based Healthcare 13, 141–146.

Poth, C., Ross, S. (2009). Meta-analysis, systematic review, or scoping review? Comparing methodologies in educational research. Paper presented at the Annual Conference of the Canadian Society for the Study of Education, Ottawa, Canada.
Rae, A., Alexander, R., McDermid, J., 2014. Fixing the cracks in the crystal ball: A maturity model for quantitative risk assessment. Reliab. Eng. Syst. Saf. 125, 67–81.
Rae, A., Provan, D., Aboelssaad, H., Alexander, R., 2020. A manifesto for Reality-based Safety Science. Saf. Sci. 126 https://doi.org/10.1016/j.ssci.2020.104654.
Raj, P.K., Lemoff, T., 2009. Risk analysis based LNG facility siting standard in NFPA 59A. J. Loss Prev. Process Ind. 22 (6), 820–829. https://doi.org/10.1016/j.jlp.2009.08.019.
Randolph, J.J., 2009. A Guide to Writing the Dissertation Literature Review. Practical Assessment, Research & Evaluation 14 (13), 1–13. https://doi.org/10.7275/b0az-8t74.
Russell, T., Reimer, T., 2018. Using Semantic Networks to Define the Quality of Arguments. Communication Theory 28 (1), 46–68. https://doi.org/10.1093/ct/qty003.
Skotnes, R.Ø., 2020. Standardization of cybersecurity for critical infrastructures - The role of sensemaking and translation. In: Olsen, O.E., Juhl, K., Lindøe, P.H., Engen, O.A. (Eds.), Standardization and Risk Governance - A Multi-Disciplinary Approach. Routledge, pp. 166–180.
Swann, G.M.P., 2000. The Economics of Standardization. Manchester Business School, University of Manchester, 11 Dec 2000.
Swann, G.M.P. (2010). The Economics of Standardization: An Update. UK Department of Business, Innovation and Skills (BIS) 27 May 2010.
Tehler, H., Abrahamsson, M., Hassel, H., Månsson, P., 2020. Standardization of disaster risk management - Challenges and opportunities. In: Olsen, O.E., Juhl, K., Lindøe, P.H., Engen, O.A. (Eds.), Standardization and Risk Governance - A Multi-Disciplinary Approach. Routledge, pp. 61–78.
Thompson, G.F., 2020. Standardization, risk dispersion, and trading. In: Olsen, O.E., Juhl, K., Lindøe, P.H., Engen, O.A. (Eds.), Standardization and Risk Governance - A Multi-Disciplinary Approach. Routledge, pp. 201–216.
Toulmin, S., 2003. [1958]). *The uses of argument* (Updated, edition ed. Cambridge University Press, Cambridge.
Tranfield, D., Denyer, D., Smart, P., 2003. Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. Br. J. Manag. 14 (3), 207–222. https://doi.org/10.1111/1467-8551.00375.
Urciuoli, L., 2018. The risk of standards proliferation – An analysis of differences between private and public transport standards. TransportationResearch Part A: Policy and Practice 116, 591–602. https://doi.org/10.1016/j.tra.2018.06.028.
Walton, D., 2006. Fundamentals of Critical Argumentation. Cambridge University Press, New York.
Walton, D., Reed, C., Macagno, F. (2008). Argumentation Schemes: Cambridge University Press.
van Eemeren, F.H., Snoeck Henkemans, A.F., 2017. Argumentation: Analysis and Evaluation, (second ed.): Routledge.
Zeng, Z., Zio, E., 2017. A classification-based framework for trustworthiness assessment of quantitative risk analysis. Saf. Sci. 99, 215–226.