Animal 17 (2023) 100863

Contents lists available at ScienceDirect

Animal
The international journal of animal biosciences

Characterising sustainability certification standards in dairy production

K. McGarr-O'Brien [a,b,*], J. Herron [a], L. Shalloo [a], I.J.M. De Boer [b], E.M. De Olde [b]

[a] Animal & Grassland Research and Innovation Centre, Teagasc, Moorepark West, Fermoy, County Cork P61 P302, Ireland
[b] Animal Production Systems Group, Wageningen University & Research, P.O. Box 338, 6700 AH Wageningen, The Netherlands

ARTICLE INFO

Article history:
Received 12 January 2023
Revised 10 May 2023
Accepted 11 May 2023
Available online 2 June 2023

Keywords:
Assessment
Farming systems
Livestock
Milk
Sustainability programmes

ABSTRACT

Despite the increasing use of private certification standards to meet the demand for sustainable dairy production, research into these standards is lacking. In this paper, we characterised sustainability certification standards currently used in dairy production. A literature search for dairy sustainability initiatives revealed one hundred-and-sixteen possible standards. In total, 19 of these were determined to qualify as 'sustainability certification standards' based on our selection criteria and were available in English or Dutch language. The standards were analysed using publicly available documents of the most recent version. The analysis included three key components: (i) general characteristics of the standard (such as the geographic origin, year founded, most recent updates), (ii) a thematic coverage analysis of the sustainability themes covered in each standard and (iii) evaluation of the inherent trade-offs within each standard utilising the opposing aspects of credibility, accessibility, and continuous improvement (the 'devil's triangle'). The comparison of general characteristics of the 19 standards revealed a wide variation in the characteristics of standards such as organisation type (i.e. nongovernmental organisations, individual dairy processor or other dairy sector actors), the number of indicators included, but also in the sustainability themes they cover, and how they balance the credibility, accessibility, and continuous improvement. The environmental pillar is most frequently and comprehensively addressed, whereas the economic pillar is least frequently and least comprehensively addressed. The 'devil's triangle' trade-off analysis revealed that credibility and accessibility, from the standard's perspective, are often transparently described and assured within the documents of standards. In contrast, continuous improvement is infrequently focused upon by standards. Overall, the variability in standards may allow farmers to choose a standard that aligns with his/her conviction or stage of development but might also create consumer or farmer mistrust in standards.

Implications

Increasingly, dairy farm sustainability standards are being deployed in response to consumer and political pressure. Research into these standards is lacking. We identified and analysed 19 dairy standards (in English or Dutch languages). Variability was discovered across their areas of sustainability focus and comprehensiveness. Environmental themes were most frequently and comprehensively addressed, as compared to social and economic themes. Continuous improvement is not focused upon by most standards. The variability observed may enable farmers to choose to participate in a standard they already align with. The variability may also, however, create mistrust in the standards by farmers or consumers.

Introduction

Dairy products provide an important source of nutrition for many in the world and contribute substantially to rural livelihoods and national economies (Liebe et al., 2020; FAO, 2022). For example, dairy is the second largest agricultural sector in the European Union, employing over two million people and accounting for 12% of total agricultural output (Bas-Defossez et al., 2019), while in the United States, the dairy industry employs over three million people and accounts for 3.5% of Gross Domestic Product (IDFA, 2021). At the same time, dairy farming is associated with multiple environmental, economic, and social sustainability challenges. First, dairy farming is a significant contributor to greenhouse gas emissions (FAO, 2013; Crippa et al., 2021), and farm management practices impact soil health, water quality, biodiversity loss, land use change, and ecosystem functions (Dumont et al., 2019; Rockström et al., 2020). Additionally, the welfare of dairy animals is an important sustainability issue, of increasing focus to consumers and therefore

* Corresponding author at: Animal & Grassland Research and Innovation Centre, Teagasc, Moorepark West, Fermoy, County Cork P61 P302, Ireland.
E-mail address: keeley.mcgarrobrien@teagasc.ie (K. McGarr-O'Brien).

https://doi.org/10.1016/j.animal.2023.100863
1751-7311/

retailers (Alonso et al., 2020; Crossley et al., 2021). Furthermore, the safety and wellbeing of farmers and rural communities, and the financial viability of their businesses, are important sustainability issues and are impacted by farm management practices (Van Calker et al., 2007; Läpple and Thorne, 2019).

There is rising consumer, industry and political pressure to evaluate and increase the sustainability of dairy production and address its challenges by, for example, reducing greenhouse gas emissions and improving animal welfare (Busch and Spiller, 2018; Schiano et al., 2020). The pressure to increase dairy sustainability is reflected in both policy- and market-based requirements aimed at dairy farming (Poppe and Koutstaal, 2020). One market-based action to increase sustainability in the food system, including the dairy sector, has been the establishment of certification standards (Hatanaka et al., 2005; Mol and Oosterveer, 2015). Certification standards often set criteria that go beyond legal norms (Rueda et al., 2017). Farmers enrol in the standard, meet its sustainability criteria demands, farm performance is assessed by an auditor, and the farm's milk is then certified by the standard. This can lead to product labels or business-to-business (**B2B**) proof of certification. Ideally, the farmer participant should see a benefit to their enrolment, whether access to new markets, maintenance of existing markets or a price premium for their milk. The use of sustainability certification standards is prominent in agriculture sectors such as coffee, cocoa, and aquaculture and their characteristics, governance, and sustainability impacts have been investigated in these sectors (Samerwong et al., 2018; Dietz and Grabs, 2022; Konefal et al., 2023). For instance, several recent studies looked at how sustainability is operationalised in standards and found imbalances in the attention to environmental, economic and social issues (Osmundsen et al., 2020; Konefal et al., 2023). As a result, standards could navigate an agricultural sector to perform well on certain sustainability themes at the expense of others. While one might argue that including as many themes as possible might be preferred to reveal trade-offs, De Olde et al. (2018) suggested that the comprehensiveness of a sustainability assessment may come at the expense of the ease of implementation. Other studies, such as by Samerwong et al. (2018) and Bush et al. (2013), looked at how standards are organised and revealed the tensions and trade-offs that exist depending on the aim, focus, administration, and compliance control methods of certification schemes. A framework called the devil's triangle was created to illuminate these trade-offs, using three opposing aspects of standard composition: credibility, accessibility, and continuous improvement (Bush et al., 2013). In this framework, credibility relates to the scientific rigour and (transparency of) stakeholder inclusion, accessibility focuses on the support offered to participants in the standard, whereas continuous improvement features how standards and participants are making improvements over time (Samerwong et al., 2018; Bush et al., 2013).

While there is an increasing focus on dairy sustainability and the number of sustainability certification standards is increasing, there is no overview of such standards, or insight into their characteristics or sustainability focus (De Olde and Busch, 2022). The objective of this paper was therefore to provide an overview of the characteristics of current dairy sustainability certification standards and analyse their contribution to integrated sustainability, and balancing of the devil's triangle aspects (credibility, accessibility, and continuous improvement). This will give a better understanding of the role standards currently play in the dairy sector, and how they operationalise sustainability. Moreover, these insights facilitate various stakeholders including standard developers, farmers, dairy organisations, and policy makers to better position the challenges and contribution of standards in the transition towards more sustainable dairy production and identify pathways to harmonise approaches internationally.

We first describe the methods used to establish existing dairy sustainability standards and their characteristics. The results and discussion sections will present the characteristics, sustainability (sub)themes addressed, and balancing of the devil's triangle trade-offs. We end with a reflection on our findings and the implications for the sustainability of dairy production.

Material and methods

Dairy sustainability standard selection process

To characterise dairy standards, we first created a database of dairy sustainability initiatives (Fig. 1). While sustainability certification can be referred to as 'schemes', 'standards', or 'labels', for this paper, we use the term 'standard'. We searched for scientific articles using Scopus and Web of Science (keywords: 'Sustainab*', 'milk', 'certification', 'scheme' or 'standard') to identify standards discussed in publications and explored the online sustainability standards databases standardsmap.org, ecolabelindex.com, and foodprint.org (keywords: 'dairy', 'milk', and 'livestock'). Finally, an internet search was performed using the keywords 'dairy sustainability standards or schemes'. These searches created a total list of 116 dairy sustainability initiatives after de-duplication (Supplementary Table S1).

Second, we applied five inclusion criteria to select relevant standards. Criteria were defined based on our research objective (e.g. focus on dairy farms (#1)) and underpinned by previous work in this field, including de Olde et al. (2016) and Konefal et al. (2023).

1. Specifically focused on dairy production at the farm-level.
2. Including internal or external audits of required criteria, resulting in 'certification' of products. Certification could include product labels or --B2B certification.
3. Encompassing more than one sustainability theme, therefore attempting to be comprehensive in their assessment of sustainability, as opposed to targeting a single sustainability theme, for example animal welfare.
4. Private or semi-private standards (e.g. privately administered with government support, but not a policy initiative).
5. Available in English or Dutch languages.

Only standards with an internal or external audit were included to ensure an actual certification process and compliance control (#2). Standards that focus on one single theme of sustainability, or only on quality assurance of the milk produced (More et al., 2021), were excluded (#3) as they are not perceived as sustainability standards. Moreover, we focused on private or semi-private standards to explore initiatives that go beyond legal standards and do not include national policies or public standards (#4). One exception to this rule was made regarding organic schemes. Organic schemes are present in many countries, often include government input or support, and often certify milk under government-established rules and/or an umbrella organisation, for example the International Federation of Organic Agriculture Movements (**IFOAM**) (IFOAM, 2022). Evaluating all the organic schemes individually would not reflect a diversity of approaches, as many are very similar. Instead, we chose to include the IFOAM standard itself, as well as two private European organic standards which include 'higher' standards than those required by EU organic regulations (i.e. the Irish Organic Association and Soil Association GB standards).

Based on these criteria, a total of 94 programmes/initiatives were excluded at this step, because they focused only on a single theme of sustainability, such as animal welfare (37 initiatives), did not include certification of products through an audit process

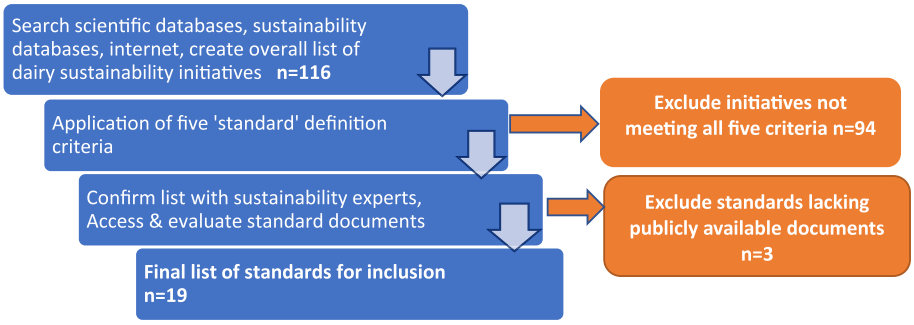

Fig. 1. Description of the steps taken to identify and compile a list of dairy sustainability certification standards to include in this paper.

(24 initiatives), represented a public policy demand (17 initiatives), or were not in the English or Dutch language (12 initiatives). The websites of the 22 remaining standards were then checked for publicly available information/documents. If limited or no information could be found, the standard organisation was contacted, and information/documents were requested. If there was no reply, or no information/documents were provided, the standard was excluded. This led to the exclusion of three further standards. A final list of 19 standards to characterise and evaluate was then confirmed. To verify the comprehensiveness of this final list of standards, the selected standards (19) were presented to two experts in sustainability programmes in the dairy sector, in the global context. They were asked if any were known to be missing. Experts confirmed the comprehensiveness of our list and saw no reason for any further inclusion or exclusion of standards.

Data collection and analysis

Using publicly available documentation from each standard website, data were collected to characterise each standard. The most recent version of the standard was used. These characteristics were divided into three categories: general characteristics, sustainability coverage and the devil's triangle framework. These categories and respective methods of analysis are described below.

General characteristics

Based on the publicly available documents, an overview of the general characteristics of each certification standard was made. The overview was made in Microsoft Excel and included the following characteristics: standard name, founding organisation and type of organisation (e.g. non-governmental organisation, dairy processor, or other dairy sector actor), country of origin, founding date, date of the most recent standard version, type and frequency of audit process, sustainability dimensions addressed (e.g. environmental, and/or social, and/or economic), and the number of indicators per standard. The audit process is the process of verifying a farmer participant is meeting standard requirements and is necessary for certification of the farm's milk. Audits can be undertaken by either an auditor internal to the standard organisation or an external auditor contracted for the job.

Sustainability coverage analysis

To analyse the sustainability subthemes addressed by each standard, the United Nations Food and Agriculture Organisation (**FAO**) Sustainability Assessment of Food and Agriculture Systems (**SAFA**) guideline was used as a reference (FAO, 2014). The FAO introduced the SAFA framework as a global tool for analysing sustainability focus across multiple food and agriculture contexts (FAO, 2014). We chose SAFA for its global applicability, and as a suitable example of a comprehensive list of sustainability subthemes used for sustainability coverage analysis of standards (Macpherson et al., 2020; Konefal et al., 2023). The SAFA guidelines distinguish four sustainability dimensions (e.g. governance, environment, economy and social), with associated themes (e.g. materials and energy), more specific subthemes (e.g. energy use), and associated indicators (e.g. percentage of energy use coming from renewable sources). We note that SAFA considers Good Governance as one of the four dimensions of sustainability for agri-food systems. However, because our evaluation of dairy sustainability standards focused on farm-level sustainability challenges only and did not deal with broader supply-chain governance issues, we chose not to include the governance dimension in our analysis. As a result, we analysed the extent to which the dairy standards covered the three sustainability dimensions, 16 themes, and 44 subthemes included in SAFA (FAO, 2014). Additionally, within this coverage analysis, we decided to present animal welfare subthemes separated from other sustainability issues. Although in SAFA, it is categorised as part of the environmental dimension, in literature, animal welfare is commonly categorised as part of the social dimension, e.g. Van der Linden et al. (2020). To avoid any conflict over which dimension should contain the animal welfare theme, we decided to list animal welfare separately in the coverage analysis.

Several recent publications have applied a coverage analysis to analyse to what extent sustainability themes are covered in standards (Sonderegger et al., 2022; Konefal et al., 2023). We carried out the coverage analysis by first creating a document with all 44 SAFA sustainability subthemes listed, divided into the 16 sustainability themes and three sustainability dimensions (Supplementary Table S2). We then accessed the certification requirements of each standard and applied the criteria (referred to in this paper as indicators) for each standard to applicable subthemes. When a standard's indicator was applicable to a particular subtheme, we noted this on the spreadsheet and considered the subtheme to be addressed by the standard. This procedure does inherently involve an interpretation of whether something is covered, or not. To ensure consistency, the same person undertook the entire coverage analysis. A distinction was noted on the spreadsheet whether the subtheme was explicitly addressed (e.g. the description within standard documents that pertain to 'animal health' exactly matched the subtheme 'animal health'), or implicitly addressed (e.g. though a standard did not use the term 'land degradation', the standard documents described how it addressed 'land degradation' by requiring soil to be covered in vegetation year-round, thus preventing degradation).

Devil's triangle analysis

In addition to outlining the sustainability focus of each standard using the SAFA framework, we investigated the trade-offs inherent

in sustainability standards. For this investigation, we introduced a definition framework based on the aspects of the devil's triangle (Bush et al., 2013). For the framework, each devil's triangle aspect- credibility, accessibility, and continuous improvement- were allocated four features (Table 1). Credibility features include the scientific rigour, inclusion of stakeholders in creating/improving standards, and transparency of scientific and stakeholder inclusion. Accessibility features address how readily participants are supported to take part in standards, as well as the transparency of standard documents. Finally, continuous improvement features include how standards keep up to date with scientific advancement, and whether participants are supported to continue improving sustainability after becoming certified in the standard (Samerwong et al., 2018; Bush et al., 2013). Each standard was evaluated, determining whether it addressed (Y) or did not address (N) each of the four features per aspect, and assigned one point for each (Y) (Supplementary Table S3). Each standard was then allocated a score (out of 4) for each of the three devil's triangle aspects; giving a highest possible total score of 12. The higher the score per aspect, the higher the standard's focus on that aspect. This method illustrates the focus of each standard across the three aspects, and therefore, the balancing of the trade-offs within the standard.

Results

The results of our analysis are outlined in this section, first with an overview of all 19 standards and their general characteristics. Next is a description of the SAFA sustainability subthemes addressed by each individual standard, as well as by the entire set of standards. Finally, an outline of the results of each standard's focus regarding trade-offs of credibility, accessibility, and continuous improvement (the devil's triangle) based on the definition framework developed for this paper.

General characteristics of standards

We found 19 dairy certification standards that met our criteria (Table 2). As expected, due to the English and Dutch language criteria, all standards evaluated in this paper originated in Europe or North America. Three distinct sub-groups of standard organisations emerged during analysis. One distinct group is non-profit organisations, also known as non-governmental organisations (**NGO**) (n = 9). Another group are standards organised by individual dairy processors or dairy co-operatives, aimed at their members, and often a prerequisite of selling milk to the processor (n = 3). Finally, there are standards organised by a range of dairy sector actors, sometimes with government support (n = 7). In this last sub-group, multiple standards began as product quality assurance schemes and have now pivoted to include environmental or other sustainability indicators within their standard (e.g. Canada ProAction, GlobalG.A.P., Red Tractor). The age of the scheme varies, the oldest being the organic standards IFOAM (1972), Irish Organic Association (**IOA**) (1973), and Soil Association (1982), which existed even before EU organic certification being introduced in 1991. The late 1990s/early 2000s saw the introduction of Demeter UK Biodynamic, Food Alliance in the USA, GlobalG.A.P., and Red Tractor UK certification standards. Notably, each of these standards have undergone many changes since they began. In fact, GlobalG.A.P. will cease to be an active livestock/dairy certification standard in 2023 (GlobalG.A.P., 2022). Origin Green and IFOAM have the least recently updated standards (2013 and 2017, respectively), eleven standards are more than five years old but have been updated over the past three years, while five standards have been newly introduced in the past five years. Some of the most recently introduced standards were previously deployed as product quality assurance (e.g. Canada ProAction, Red Tractor) or animal welfare (e.g. Beter Leven) initiatives, but have since incorporated more sustainability themes, making them fit the qualifying criteria for inclusion in this paper as certification standards. This illustrates the evolving nature of some standards as they adapt to incorporate sustainability into their programmes, whereas others (e.g. Food Alliance, Leaf Marque) have had a sustainability focus since first being introduced.

The number of indicators (criteria required for certification) used by each standard varies from a low of 18 (Caring Dairy) to a high of 436 (IOA), with an average of 141 indicators per standard. The number of indicators per standard is complicated in some cases, however. Some standards require participants to first be members of underlying quality assurance or organic programmes which are not included in the indicators listed for the standard itself. For example, Leaf Marque requires participating in Red Tractor as a baseline, and Caring Dairy requires participation in the Cono quality assurance baseline. Moreover, standards demand adhering to criteria determined by law in the jurisdiction of the standard. For this reason, the total criteria requirements for participating in each standard are not always obvious only from standard documents. In addition to variability in the number of indicators, seven standards in this paper include indicators that are recommended rather than required. Some state the recommended indicators are a method to enable context-specificity and to allow farmer participants flexibility, others state recommended indicators assist in continuous improvement, for farmers to go 'above and beyond' where possible.

To certify standard participants, organisations use verification systems consisting of audits by either their own internal auditors or external auditors contracted for this purpose. Of the 19 standards evaluated, audits are performed by external auditors for 12

Table 1
Devil's triangle framework aspects and features per aspect, applied to analyse dairy sustainability standards.

Devil's triangle aspect	Feature definitions per aspect			
	Feature 1	Feature 2	Feature 3	Feature 4
Credibility	Inclusion of farmer/farm representative stakeholder groups to create and/or update standard.	Standard states that scientists/ researchers advise or participate in creating/updating standard criteria.	Based on available documents, standard is transparent about its scientific background.	Based on available documents, standard is transparent about its verification system.
Accessibility	Support is available to prospective and current participants (e.g. transition fund, pre-assessment, translations, advisory, etc).	Standard fees: Is a sliding scale based on farm size or volume of sales available?	Availability of farmer support to those participating (peer groups, information exchange events, online platform, etc).	Standard documents are transparent and readily accessible to all.
Continuous Improvement	Multiple standard levels (e.g. gold, silver, bronze)? and/or multiple levels within the standard (e.g. 1,2,3,4 within criteria).	Standard incorporates a way for farmers to continue improving after initial certification (within same standard).	There are stated regularly scheduled updates to standards.	Standards are transparent about methods for updating and incorporating new scientific information.

Table 2
General characteristics of 19 evaluated dairy sustainability certification standards.

Standard Name	Organisation Name & Type (NGO/ Processor/Dairy Sector Actor)	Origin Country	Origin year, most recent version	Audit type & frequency	Sustainability dimensions addressed[2]	Number of indicators
Better For/Beter Voor	Albert Heijn/Royal Aware (Processor)	Netherlands	2017, 2020	External, annual	All 3	48
AWA by AGW EU	A Greener World (NGO)	United States	2014, 2021	External, annual	Environment	257[1]
AGA Grassfed	American Grassfed Association (Sector Actor)	United States	2009, 2022	External, 12–15 months	Environment	78
Arlagården	Arla Dairy Cooperative (Processor)	Denmark/ Sweden	2003, 2021	External, 3 years	All 3	117
Better Life/Beter Leven	Better Life Label/Dutch Society for the Protection of Animals (NGO)	Netherlands	2021, 2021	External, annual	All 3	213
Canada ProAction	Dairy Farmers of Canada (Sector Actor)	Canada	2017, 2021	Internal, 2 years	Environment, economic	82[1]
Caring Dairy	Cono Cheesemakers (Processor)	Netherlands	2004, 2022	External, annual	Environment, economic	18
Certified Naturally Grown (CNG)	CNG (NGO)	United States	2002, unknown	PGS peer-to-peer, annual	Environment	67
Demeter Biodynamic UK	BFDI (NGO)	Germany	1999, 2021	Internal, annual	All 3	193
Food Alliance	Food Alliance (NGO)	United States	1997, 2022	External, 3 years	Environment, social	47
GlobalG.A.P.	FoodPlus LLC (Sector Actor)	Germany	1997, 2019	External, annual	All 3	217[1]
IFOAM	IFOAM (NGO)	Germany	1972, 2014–17	Up to standard organisations	All 3	114[1]
Irish Organic Association (IOA)	IOA (Sector Actor)	Ireland	1982, 2022	Internal, annual	All 3	436[1]
Leaf Marque	LEAF (NGO)	United Kingdom	2012, 2020	External, annual	All 3	100
Origin Green	Bord Bia (Sector Actor)	Ireland	2013, 2013	External, 18 months	All 3	170[1]
On the Way to PlanetProof	SMK (NGO)	Netherlands	2018, 2022	External, 2 years	All 3	54
Pasture for Life Association (PFLA)	PFLA cic (Sector Actor)	United Kingdom	2018, 2018	External, annual	Environment, economic	154
Red Tractor	Red Tractor Assurance (Sector Actor)	United Kingdom	2000, 2021	Internal, 18 months	All 3	161
Soil Association Organic	Soil Association Charity (NGO)	United Kingdom	1973, 2022	Internal, annual	Environment, social	147[1]

Abbreviations: NGO = non-governmental organisation; AWA by AGW = Animal welfare approved by A Greener World; AGA = American Grassfed Association; PGS = Participatory Guarantee System; BFDI = Biodynamic Federation Demeter International; G.A.P = Good Agricultural Practices; IFOAM = International Federation of Organic Agriculture Movements; LEAF = Linking Environment and Farming' SMK = Stichting Milieukeur; cic = community interest company.
[1] In addition to required indicators, recommended indicators are also present in standard.
[2] Which sustainability dimensions are addressed by the standard's indicators: environmental and/or social and/or economic (if all 3 dimensions are addressed, 'all 3' is entered).

of the standards. A further five standards certify their product with audits performed by internal auditors. Two standards have unique audit arrangements. Firstly, IFOAM, an organic 'umbrella' organisation, provides its standard to be adopted by other certification organisations, such as the Bio Suisse Standards from Switzerland, who will then determine whether to provide internal or external audits. Secondly, the Certified Naturally Grown standard which relies upon a peer-review audit called participatory guarantee system (**PGS**) for its certification. This style of peer-to-peer inspection is aimed mainly at small-scale farmers, with the aims of local knowledge sharing and networking between farmers, while certifying farming practices. The PGS process, in general, is in use in approximately 242 agricultural initiatives worldwide (IFOAM, 2022).

Sustainable assessment of food and agriculture systems coverage analysis

Sustainable assessment of food and agriculture systems subthemes addressed per dimension for individual standards

The results of the coverage analysis (i.e. the number of subthemes addressed by each of the 19 individual standards), by sustainability dimension are shown in Fig. 2. This analysis shows the variability in the number of SAFA subthemes addressed per standard, with a low of six subthemes (AGA by AGW, Certified Naturally Grown, Red Tractor), and a high of 20 subthemes (IFOAM). The average number of subthemes addressed per standard was 12, out of a total of 44 possible subthemes. All standards address all animal welfare and some environmental subthemes, four standards do not address any economic subthemes, eight standards do not address any social subthemes, and three standards do not address either social or economic subthemes. Leaf Marque and Soil Association standards addressed the highest number of environmental subthemes, with Leaf Marque addressing all 12, and Soil Association Organic addressing 10 out of 12 subthemes. Three standards, Demeter Biodynamic, Food Alliance, and IFOAM Organic, had a notably higher focus on social sustainability than other standards, addressing eight, six, and nine out of 16 possible subthemes, respectively. Caring Dairy and Origin Green addressed the highest number of economic subthemes (four and three, respectively); however, this was still a minority of the 14 possible economic subthemes.

Total sustainable assessment of food and agriculture systems subtheme coverage per sustainability dimension across all 19 standards

The coverage of subthemes by all standards are provided in Figs. 3–5, by sustainability dimension. Overall, the environmental

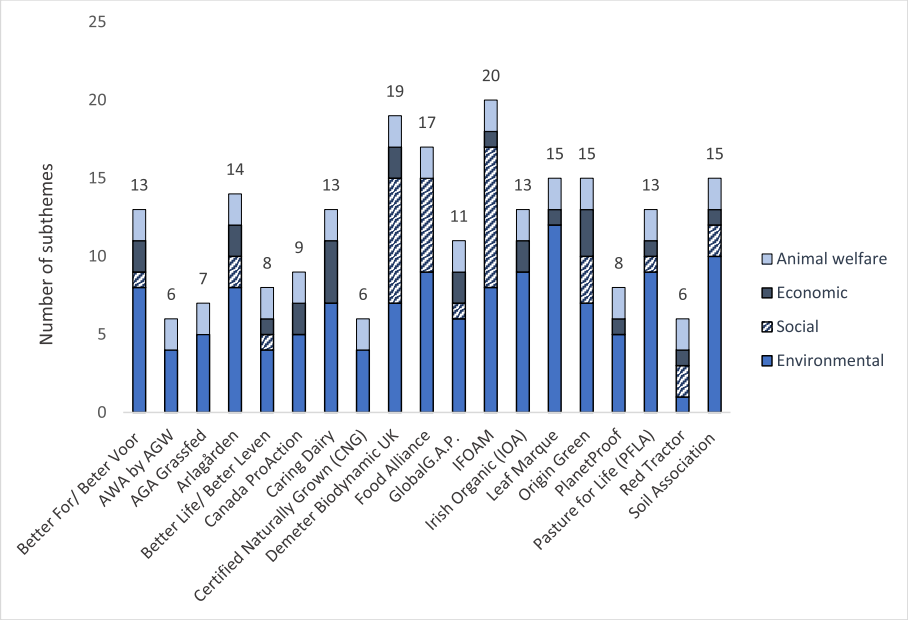

Fig. 2. Sustainable Assessment of Food and Agriculture Systems (SAFA) sustainability subthemes addressed per dairy sustainability standard. Maximum number of subthemes is 44 (12 environmental, 16 social, 14 economic, 2 animal welfare). Abbreviations: Please refer to Table 2 for abbreviation definitions.

and animal welfare dimension was addressed most frequently and comprehensively, the social dimension was the second most frequently addressed, and the economic dimension the least frequently and least comprehensively addressed. The two animal welfare subthemes from SAFA, animal health and freedom from stress, were both addressed by all 19 standards (Fig. 3). Each of the 12 environmental subthemes was addressed by multiple standards, with water quality, soil quality, and both ecosystem and species diversity most covered. The least covered environmental subthemes were water withdrawal and waste reduction & disposal (Fig. 3). In the social sustainability dimension, workplace health & safety was the most frequently addressed subtheme, while six of the 16 subthemes were not addressed by any standard (Fig. 4). In the economic sustainability dimension, community investment, food safety, food quality, and local procurement were addressed, whereas 10 out of 14 subthemes were not addressed by any standard (Fig. 5). The safety and quality-focused subthemes within social and economic sustainability are the most frequently addressed. Six standards address the subtheme local procurement, and in all cases, these were requirements to grow or purchase local forage or local grain.

Devil's triangle analysis

Balancing of trade-offs

We evaluated how each certification standard balances the devil's triangle as described in the method section. Fig. 6 illustrates the outcomes with the scores per aspect per standard. Eleven standards focus solely on credibility and accessibility, scoring 0 out of 4 for continuous improvement. On average, standards scored high-

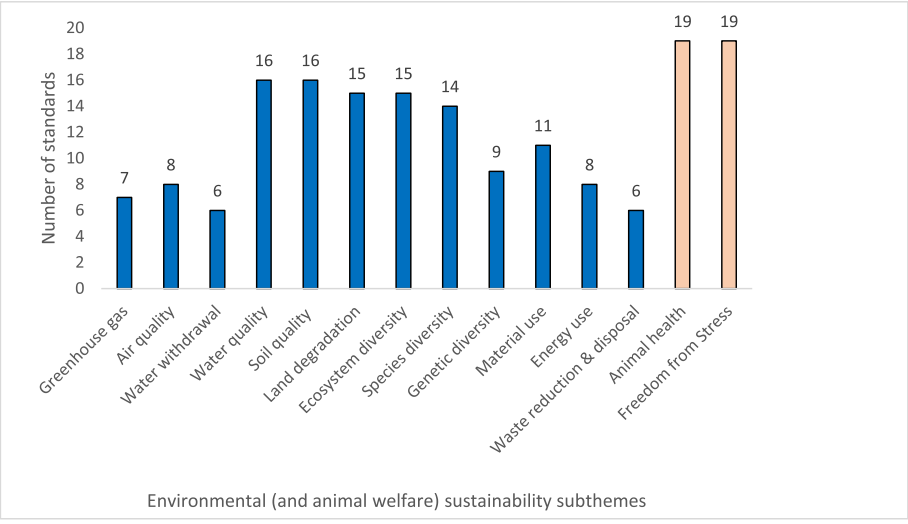

Fig. 3. Total dairy standard subtheme coverage of the Sustainable Assessment of Food and Agriculture Systems (SAFA) environmental and animal welfare sustainability dimensions. Maximum number of subthemes = 19. Note: Animal welfare subthemes 'animal health' and 'freedom from stress' indicated by distinct colour (refer to material and methods for more information).

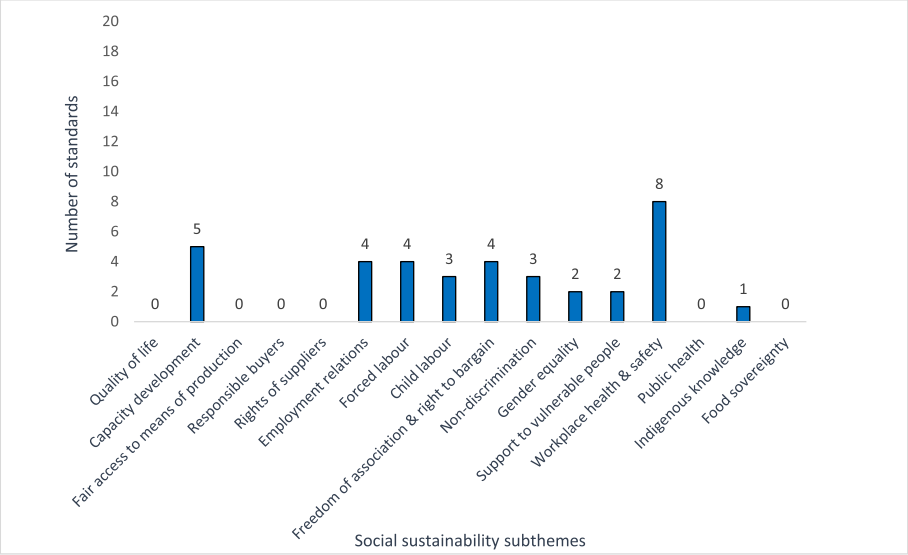

Fig. 4. Total dairy standard subtheme coverage of the Sustainable Assessment of Food and Agriculture Systems (SAFA) social sustainability dimension. Maximum number of subthemes = 19.

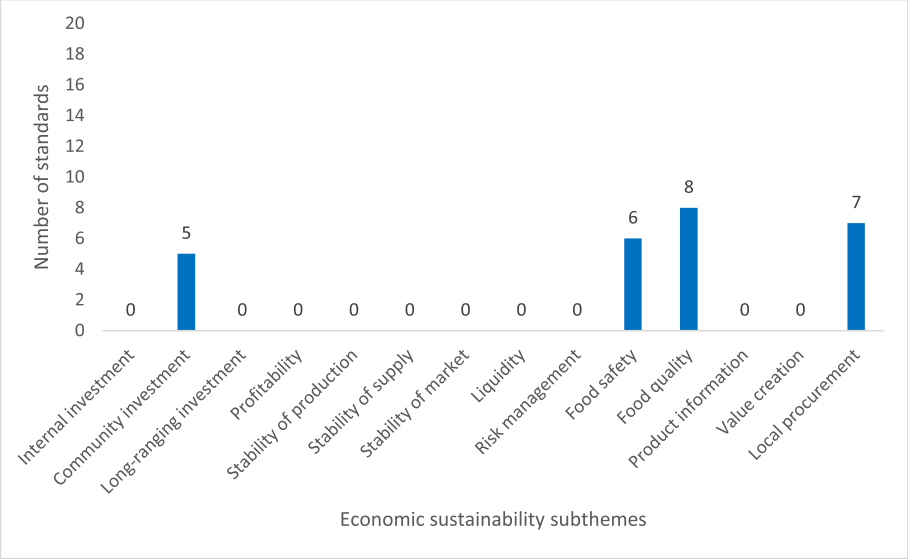

Fig. 5. Total dairy standard subtheme coverage of the Sustainable Assessment of Food and Agriculture Systems (SAFA) economic sustainability dimension. Maximum number of subthemes = 19.

est on their credibility with 3.6/4, accessibility scored second highest with 3.1/4, and continuous improvement had the lowest score with 0.9/4. Three standards, Caring Dairy, Food Alliance, and Leaf Marque showed the most balance between the three aspects, each having a total score of 10 or above out of a possible 12. Leaf Marque was the only standard to score 12 out of 12. Notably, these standards (Caring Dairy, Food Alliance, and Leaf Marque) all state within their documents that an important part of their strategy is to enable continuous improvement of farm practices and of the standards themselves.

Scores per devil's triangle aspects: credibility, accessibility, continuous improvement

Features for both credibility and accessibility include transparency, of verification system (credibility) and of documents (accessibility). Unsurprisingly, most standards scored points for these features, since the standards had sufficiently transparent access to their documents for inclusion within this paper. Most standards state they incorporate farmers as well as scientific researchers within their committees on creating and updating standards, leading to higher scores in credibility. Seventeen of the standards stated they involve multiple stakeholder groups, including farmers or farmer representatives, in creating and updating standards. Descriptions of governance style commonly described one or multiple 'board of experts' or 'technical advisory committee (TAC)' who advise on standards. However, transparency surrounding who these experts are and how they were appointed was variable.

Features defining accessibility included support for participants, both organisationally and financially. Most standards offer some style of support to farmers to participate, but there is variation in the type of support offered, for example online information hubs and peer-to-peer learning opportunities. A financial access

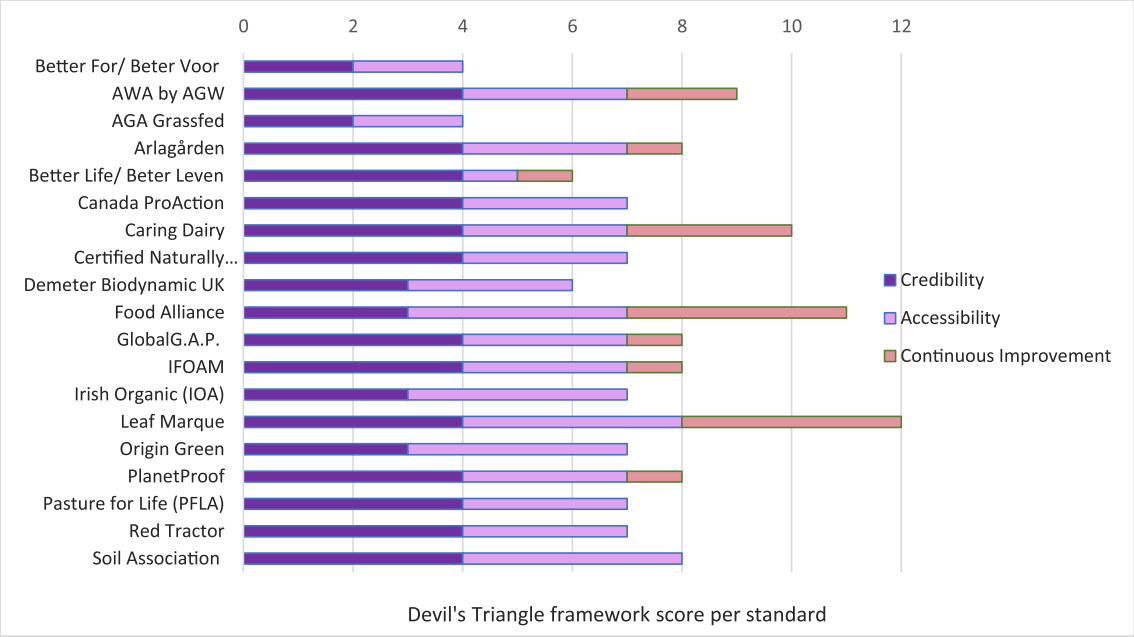

Fig. 6. Score per dairy standard, per aspect of credibility, accessibility, and continuous improvement. Highest possible score per aspect = 4. Highest maximum possible score = 12. Abbreviations: Please refer to Table 2 for abbreviation definitions.

feature was the availability of a sliding fee scale to enable farmers more equitable access to participate in the standard regardless of farm size. Nearly, all standards offer a sliding scale for annual membership, or have no fees at all, in the case of processor standards required for selling milk. However, there remains variability in affordability of standards, as sliding scales for membership are not the only fees. For example, some standards require farmers to cover the costs of audits, as well as pay a one-time fixed registration fee. Additionally, participating in a standard could require changes in labour (e.g. required training) and farm management (e.g. lower stocking density) or investments in equipment and technology (e.g. emission reducing housing systems, monitoring systems or precisions fertilisation machinery) (Meemken, 2020).

Continuous improvement features included whether standards are regularly updated, and whether there is transparency surrounding how improvements are made, neither of these were met by many standards (five and six standards, respectively). Three standards gained points for the feature of farmers being able to continue improving after initial certification (Arlagården, Food Alliance, and Leaf Marque). This happened in all cases with voluntary self-assessments or goal reviews and updates. Another feature of continuous improvement is whether multiple levels exist within the standard, only four standards meet this definition (Beter Leven, Caring Dairy, Food Alliance, and Leaf Marque).

Discussion

The aim of this study was to identify, characterise, and analyse dairy sustainability certification standards. We found 19 standards, and by analysing their characteristics, revealed heterogeneity in characteristics, such as the organisation type (i.e. NGO, individual dairy processor or other dairy sector actors), the sustainability themes covered, the number of indicators included, and how standards balance credibility, accessibility, and continuous improvement.

Indicators are a key component of standards, as they represent the requirements of the standard and are therefore the means through which standards demonstrate the sustainability credentials of participants (Osmundsen et al., 2020). The variability in indicators utilised per standard can lead to concerns in the validity of standards and may erode public trust in standards (De Olde et al., 2017). Our analysis found variability between standards in the sustainability dimension the indicators most frequently address, while the environmental dimension was addressed by the highest number of indicators, there was no consistency across standards, and this carried through to the social and economic dimensions as well. Also, the number of indicators used per standard varied, ranging from 18 to 436 indicators. These findings align with previous research that revealed a bias towards environmental sustainability over social or economic sustainability and that the choice of indicators varies widely between standards (Konefal et al., 2023; Osmundsen et al., 2020; De Olde et al., 2017).

In the standards we analysed, not all indicators had to be met by producers. In seven standards, both recommended and required indicators were included. The flexibility of recommended indicators may be welcomed by farmer participants; however, verification is less transparent in standards incorporating recommended indicators. If a participant does not see some benefit to implementing a recommended indicator, there is a lack of incentive to adhere to a recommendation (as opposed to a requirement). Conversely, if the recommended indicator is of benefit (for example economic benefit) to the farmer, it is reasonable to assume they would have implemented it regardless of whether they are a standard participant; although, adoption of sustainable practices varies depending on the farmers' motivation and attitude towards sustainability (Luhmann et al., 2016). There are differences, therefore, in conclusions that can be drawn about sustainability changes resulting from standard participation where recommended indicators are used.

Differences in indicator type and quantity between standards illustrate that the demand standards make on participating farmers is variable. Further differences were noted between standard verification type (e.g. twelve standards used external audits, five internal audits, and two used other approaches) and the requirements to comply with verification audits. For example, the Animal Welfare Approved by A Greener World (AWA by AGW) standard states that 'any practice not covered in the standard is permissible

unless specifically excluded', as they acknowledge the context-specificity of animal husbandry practices and farmer knowledge. In practice, this means the AWA by AGW standard has a range of required and recommended indicators and places a strong emphasis on the role of auditors in confirming that participants have met the standard requirements. In contrast, the Demeter Biodynamic standard states its list of requirements should be seen as a 'positive list', e.g. if something is not mentioned on the list of requirements it is not allowed unless permission is asked for. From these differences, we derive that the costs for participating farmers (e.g. time, financial, resource costs) will also vary.

Given the large variety in the number of indicators included in standards, a follow-up study to explore the different types of indicators (i.e. target, practice or performance (FAO, 2014)) is relevant to provide insight into the contribution of farming practices to sustainability goals. A detailed categorising of the types of indicators used by standards would more accurately reflect the potential sustainability impacts of the standard, as well as differences between standards.

Sustainability coverage analysis

Dairy sustainability certification standards are market-oriented; therefore, consumer trust in their credibility is important to enable them to continue (More et al., 2021). This requires transparency as well as consumer assurance regarding the effectiveness of standards to contribute to aims they find important. If the aims of multiple dairy standards – as represented by their sustainability focus – are variable, this could create a mismatch in what is covered in a standard and what is expected, which in turn may affect consumer trust (Hoogland et al., 2007).

The sustainability focus of each standard was established by the sustainability dimensions and SAFA subthemes it addressed, as revealed by our sustainability coverage analysis. This analysis illustrated a large variation between standards' areas of sustainability focus both in terms of the number of subthemes covered (ranging from 6 to 20, out of a total of 44), and in the subthemes covered. Subthemes within the environmental dimension are more often covered in standards, whereas several economic and social subthemes were covered by none of the standards. Within social and economic sustainability, safety and quality-focused subthemes were most frequently addressed, which can be explained by the previously mentioned evolution of some standards from quality assurance into more comprehensive sustainability standards.

In addition to variability of sustainability dimensions and subthemes addressed per standard, each standard operationalises sustainability in a different way. For example, in the soil quality subtheme, the Leaf Marque standard requires measuring soil quality by visual assessment, earthworm count, and a soil test, as well as requiring measures be taken to build up soil organic matter. In contrast, other standards (e.g. Origin Green) require or recommend soil testing at specified intervals (e.g. every three years). In both examples, standards are addressing 'soil quality' as a sustainability subtheme but this could lead to different (or no) farm management change. Additionally, in the latter example, no corrective action is required after soil testing, so the final step of improving soil quality is not guaranteed. This aligns with prior research on sustainability assessment which found that undertaking the assessment process is not a guarantee of a sustainable transformation (Alrøe and NOE, 2016). While SAFA coverage analysis can help identify the intention of standards, the overall sustainability impact cannot be established without examining the indicators, outcomes, and perspective of standard organisers and farmer participants in more detail, as for instance carried out for other sectors by Lamarque and Lambin (2015) and Rubio-Jovel (2022).

Balancing sustainability standard trade-offs

The development of certification standards relies upon striking a balance between trade-offs, including varying requirements of multiple stakeholders such as standard organisers, participating farmers, consumers, livestock, and the environment. This balance has been conceptualised in several ways. De Olde et al. (2018) reveal a 'comprehensiveness continuum', which implies that the more comprehensively sustainability (sub)themes are addressed within a standard, the more difficult it is to participate. Binder et al. (2010) also address trade-offs being inherent to the nature of sustainability assessments. This is partly because in practice, standards must be selective about the sustainability (sub)themes they address. Consequently, they focus on specific (sub)themes, which can lead to ignoring the adverse effects on others. This lack of an integrative approach, as outlined by Alrøe and NOE (2016), relates not only to the sustainability focus but also to the organisation of the standard. The 'devil's triangle' framework (Bush et al., 2013), which illustrates standard trade-offs, identified credibility, accessibility and continuous improvement as three aspects of standard organisation that are often at odds with each other (Samerwong et al., 2018).

The devil's triangle analysis of this paper found that standards balance the trade-offs between credibility, accessibility, and continuous improvement in different ways. It is important to point out that the elements and resulting scores are derived only from information in the documents provided by each standard and are only providing insight into the standard at that point of time. It is not possible from these documents to evaluate (or 'score') the standards based on the actual farm performance or the perspective of participating farmers on standard credibility, accessibility, or continuous improvement.

Across all standards, credibility, illustrated by features of transparency and multi-stakeholder and scientific expert input in standard design, was most focused on. Due to standard organisations investing significant resources to develop and deploy standards, this finding is reasonable. Accessibility features include how well-supported farmers are to become standard participants. Accessibility was nearly as often focused upon as credibility. This also seems reasonable, as standard organisers (whose perspective is being reflected upon in this paper) would want to involve as many farmers as possible. Continuous improvement, which focuses on (the transparency of) updates to the standard, was by far the least focused on aspect of the devil's triangle. It is possible this information is known internally to the standard organisers, but not transparent within publicly accessible standard documents. Moreover, introducing changes or updates in standards requires an investment of resources (i.e. negotiating changes within and outside the standard organisation, facilitating feedback on proposal, updating documentation and audit procedures, informing and/or training participants). This may explain why the process of continuous improvement is less of a focus. Continuous improvement is, however, a critical aspect of updating standards to incorporate new scientific research and ensure ongoing sustainable development (Konefal et al., 2023). Interestingly, our analysis of continuous improvement revealed that the three standards that scored highest on continuous improvement, Caring Dairy (score of 3), Food Alliance (score of 4), Leaf Marque (score of 4), were all established and intended as 'holistic' sustainability certification standards. This contrasts with other standards that began as single-issue standards (e.g. grassfed) or product quality assurance standards and have evolved to include aspects of sustainability, all of which scored 0 or 1 in continuous improvement.

Implications for dairy farmers

Dairy farmers, as those faced with the sustainability criteria defined by standards, will be confronted with the impact of the

heterogeneity identified in this paper. First, by defining criteria for producers to meet, standards determine what is thereby perceived as sustainable, and what is not (Osmundsen et al., 2020). Differences among standards might then result in differences in interpretation of, for example, which (sub)themes are relevant, and which are not. These differences in interpretation might affect farmer's trust in whether these standards help them farm more sustainably. Differences in standards, moreover, can result in competition, and, for instance, choosing those requirements that align with your convictions or are easier to fulfil. In a way, this could offer opportunities for producers for continuous improvement, and start with an 'easier' standard and move up. Nevertheless, it might also contribute to consumer's mistrust in the value of standards for the transition towards sustainable farming.

Our findings, furthermore, revealed an inherent conflict of standards. On the one hand, standards require participating farmers to work towards an end-point for certification, whereas, on the other hand, sustainable development in itself is not a fixed state but rather an ongoing process of change also influenced by the perspectives and values of those involved (Bell and Morse, 2008). This finding aligns with the analysis of Amundsen (2022), who describes how the inflexible nature of certification standards challenges the ability for standards to continually improve. Standards focus on applying certain farm practices or management measures, rather than on the actual improvement to increase sustainable outcomes. Amundsen (2022) argues that without incorporating continuous improvement, standards become 'check-box exercises', and may not actually impact sustainability. A trend to move away from a 'check-box' approach towards measuring impact on farm, using performance-based indicators, has been recently described by Hatanaka et al. (2022) in the case of US crop production. Performance-based approaches are, however, according to Alrøe et al. (2017), in contrast with so-called value-based approaches that focus on the communication of sustainability values to support coordinated and cooperative action. Performance-based approaches risk not paying sufficient attention to the context and the values of stakeholders involved, and steer towards what is easy to measure. Both approaches present their own strengths and weaknesses, rather than choosing one or the other, a reflexive approach is needed in which different perspectives, rationales and values are embraced (Alrøe et al., 2017).

Future direction for dairy sustainability standards

The next step to further understanding dairy sustainability certification standards should move beyond this paper's characterisation which is based solely upon standard documents and undertake more detailed investigation. For instance, thorough indicator analysis will help gain insight into the potential sustainability impact resulting from standard participation. Additionally, analysing standard attributes from the perspectives of both standard organisation representatives and farmer participants will offer further insight into standards and their sustainability credentials. Furthermore, this paper focused on a specific set of standards, those that are available in English or Dutch language. To analyse the characteristics, similarities and differences in standards between geographical regions, extending the scope to include standards used around the globe would be of interest. Moreover, studying the development of standards over time, in terms of governance structures, compliance methods and sustainability focus areas, could be interesting to reveal trends.

To address the heterogeneity identified in this paper, one approach would be to try to harmonise approaches (Poppe and Koutstaal, 2020). Poppe and Koutstaal proposed linking public sustainability goals with private certification standards by identifying key performance indicators on issues such as climate, biodiversity and animal welfare, and evaluating the performance of farmers. Farmers could thereby receive payments from both public and private schemes. Such action may require private standards to undergo regulatory oversight, or to reach sector-wide, international consensus on a definition of sustainability which is then accepted by regulators (Poppe and Koutstaal, 2020). These adjustments would signify a major change to private standards. Future research investigating existing or potential alignment of private standards and public sustainability goals could contribute to this policy discussion.

Conclusion

This review of 19 standards for dairy farming revealed a wide variation in governance and verification processes of standards, as well as demands each standard places on farmer participants. This variability might allow farmers to choose a standard that aligns with his/her conviction or stage of development, but could also potentially create consumer or farmer mistrust in standards. The 'devil's triangle' trade-off analysis, moreover, revealed that credibility and accessibility, from the standard's perspective, are often transparently described and assured within the documents of standards, whereas continuous improvement is infrequently focused upon. Yet, given the importance of a sustainable development of the dairy sector, attention to continuous improvement is essential. This, however, may require moving away from a traditional approach to standards through checking boxes towards more inclusive approaches that focus on the capacity development of the farmer to facilitate the sustainable development of the farm in its unique context.

Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.animal.2023.100863.

Ethics approval

Not applicable.

Data and model availability statement

None of the data were deposited in an official repository. Information can be made available from the authors upon request.

Author ORCIDS

McGarr-O'Brien, K.: https://orcid.org/0000-0003-4505-198X.
Herron, J.: https://orcid.org/0000-0001-6704-994X.
Shalloo, L.: https://orcid.org/0000-0003-1714-672X.
De Boer, I.J.M.: https://orcid.org/0000-0002-0675-7528.
De Olde, E.M.: https://orcid.org/0000-0002-5520-7514.

Author contributions

Keeley McGarr-O'Brien: Conceptualisation, Methodology, Investigation, Writing-Original draft preparation, reviewing and editing. **Jonathan Herron:** Supervision, Writing-reviewing and editing. **Laurence Shalloo:** Conceptualisation, Writing-reviewing and editing. **Imke De Boer:** Supervision, Conceptualisation, Methodology, Writing-reviewing and editing. **Evelien De Olde:** Supervision, Conceptualisation, Methodology, Writing-reviewing and editing.

Declaration of interest

None.

Acknowledgements

None.

Financial support statement

The authors acknowledge a research grant from Science Foundation Ireland and the Department of Agriculture, Food and Marine on behalf of the Government of Ireland under the Grant 16/RC/3835 (VistaMilk).

References

Alonso, M.E., González-Montaña, J.R., Lomillos, J.M., 2020. Consumers' concerns and perceptions of farm animal welfare. Animals 10, 385. https://doi.org/10.3390/ani10030385.

Alrøe, H.F., Noe, E., 2016. Sustainability assessment and complementarity. Ecology and Society 21, 30. https://doi.org/10.5751/es-08220-210130.

Alrøe, H., Sautier, M., Legun, K., Whitehead, J., Noe, E., Moller, H., Manhire, J., 2017. Performance versus values in sustainability transformation of food systems. Sustainability 9, 332. https://doi.org/10.3390/su9030332.

Amundsen, V.S., 2022. From checking boxes to actual improvement: A new take on sustainability certification. Aquaculture 548, 737672. https://doi.org/10.1016/j.aquaculture.2021.737672.

Bas-Defossez, F., A.B., Lorant A, Kollenda E,, 2019. A Vision for the Future of the European Dairy Industry. Institute for European Environmental Policy IEEP, Brussels/London, BE/UK.

Bell, S., Morse, S., 2008. Sustainability Indicators: Measuring the Immeasurable? Earthscan, London, UK.

Binder, C.R., Feola, G., Steinberger, J.K., 2010. Considering the normative, systemic and procedural dimensions in indicator-based sustainability assessments in agriculture. Environmental Impact Assessment Review 30, 71–81. https://doi.org/10.1016/j.eiar.2009.06.002.

Busch, G., Spiller, A., 2018. Consumer acceptance of livestock farming around the globe. Animal Frontiers 8, 1–3. https://doi.org/10.1093/af/vfx005.

Bush, S.R., Toonen, H., Oosterveer, P., Mol, A.P.J., 2013. The 'devils triangle' of MSC certification: Balancing credibility, accessibility and continuous improvement. Marine Policy 37, 288–293. https://doi.org/10.1016/j.marpol.2012.05.011.

Crippa, M., Solazzo, E., Guizzardi, D., Monforti-Ferrario, F., Tubiello, F.N., Leip, A., 2021. Food systems are responsible for a third of global anthropogenic GHG emissions. Nature Food 2, 198–209. https://doi.org/10.1038/s43016-021-00225-9.

Crossley, R.E., Bokkers, E.A.M., Browne, N., Sugrue, K., Kennedy, E., De Boer, I.J.M., Conneely, M., 2021. Assessing dairy cow welfare during the grazing and housing periods on spring-calving, pasture-based dairy farms. Journal of Animal Science 99, skab093. https://doi.org/10.1093/jas/skab093.

De Olde, E.M., Busch, G., 2022. Opinion paper: Role of standards in the transformation of the livestock sector. Animal 16, 100588. https://doi.org/10.1016/j.animal.2022.100588.

De Olde, E.M., Oudshoorn, F.W., Sørensen, C.A.G., Bokkers, E.A.M., de Boer, I.J.M., 2016. Assessing sustainability at farm-level: Lessons learned from a comparison of tools in practice. Ecological Indicators 66, 391–404. https://doi.org/10.1016/j.ecolind.2016.01.047.

De Olde, E.M., Moller, H., Marchand, F., McDowell, R.W., Macleod, C.J., Sautier, M., Halloy, S., Barber, A., Benge, J., Bockstaller, C., Bokkers, E.A.M., De Boer, I.J.M., Legun, K.A., Le Quellec, I., Merfield, C., Oudshoorn, F.W., Reid, J., Schader, C., Szymanski, E., Sørensen, C.A.G., Whitehead, J., Manhire, J., 2017. When experts disagree: the need to rethink indicator selection for assessing sustainability of agriculture. Environment, Development and Sustainability 19, 1327–1342. https://doi.org/10.1007/s10668-016-9803-x.

De Olde, E.M., Sautier, M., Whitehead, J., 2018. Comprehensiveness or implementation: Challenges in translating farm-level sustainability assessments into action for sustainable development. Ecological Indicators 85, 1107–1112. https://doi.org/10.1016/j.ecolind.2017.11.058.

Dietz, T., Grabs, J., 2022. Additionality and Implementation Gaps in Voluntary Sustainability Standards. New Political Economy 27, 203–224. https://doi.org/10.1080/13563467.2021.1881473.

Dumont, B., Ryschawy, J., Duru, M., Benoit, M., Chatellier, V., Delaby, L., Donnars, C., Dupraz, P., Lemauviel-Lavenant, S., Méda, B., Vollet, D., Sabatier, R., 2019. Review: Associations among goods, impacts and ecosystem services provided by livestock farming. Animal 13, 1773–1784. https://doi.org/10.1017/s1751731118002586.

FAO, 2013. Tackling climate change through livestock: A global assessment of emissions and mitigation opportunities. Retrieved on 15 August 2022 from http://www.fao.org/3/i3437e/i3437e.pdf.

FAO, 2014. Sustainability assessment of food and agriculture systems (SAFA) guidelines. Version 3. Food and Agricultural Organization (FAO) Rome, Italy.

FAO, 2022. Gateway to dairy and dairy products. Retrieved on 15 June 2022 from https://www.fao.org/dairy-production-products/products/en/.

GlobalG.A.P., 2022. GlobalG.A.P. Livestock Standard Announcement. Retrieved on 15 October 2022 from https://www.globalgap.org/export/sites/default/.content/.galleries/Documents_Other/220614_GLOBALG.A.P._Livestock_Standard_Announcement.pdf.

Hatanaka, M., Bain, C., Busch, L., 2005. Third-party certification in the global agrifood system. Food Policy 30, 354–369. https://doi.org/10.1016/j.foodpol.2005.05.006.

Hatanaka, M., Konefal, J., Strube, J., Glenna, L., Conner, D., 2022. Data-driven sustainability: metrics, digital technologies, and governance in food and agriculture. Rural Sociology 87, 206–230. https://doi.org/10.1111/ruso.12415.

Hoogland, C.T., de Boer, J., Boersema, J.J., 2007. Food and sustainability: Do consumers recognize, understand and value on-package information on production standards? Appetite 49, 47–57. https://doi.org/10.1016/j.appet.2006.11.009.

IDFA, 2021. United States Dairy Industry's Economic Impact Totals $753 Billion. Retrieved on 1 June 2022 from https://www.idfa.org/news/u-s-dairy-industrys-economic-impact-totals-753-billion.

IFOAM, 2022. Participatory Guarantee Systems (PGS). Retrieved on 20 November 2022 from https://www.ifoam.bio/our-work/how/standards-certification/participatory-guarantee-systems.

Konefal, J., de Olde, E.M., Hatanaka, M., Oosterveer, P.J.M., 2023. Signs of agricultural sustainability: A global assessment of sustainability governance initiatives and their indicators in crop farming. Agricultural Systems 208, 103658. https://doi.org/10.1016/j.agsy.2023.103658.

Lamarque, P., Lambin, E.F., 2015. The effectiveness of marked-based instruments to foster the conservation of extensive land use: The case of Geographical Indications in the French Alps. Land Use Policy 42, 706–717. https://doi.org/10.1016/j.landusepol.2014.10.009.

Läpple, D., Thorne, F., 2019. The role of innovation in farm economic sustainability: generalised propensity score evidence from Irish dairy farms. Journal of Agricultural Economics 70, 178–197. https://doi.org/10.1111/1477-9552.12282.

Liebe, D.L., Hall, M.B., White, R.R., 2020. Contributions of dairy products to environmental impacts and nutritional supplies from United States agriculture. Journal of Dairy Science 103, 10867–10881. https://doi.org/10.3168/jds.2020-18570.

Luhmann, H., Schaper, C., Theuvsen, L., 2016. Future-oriented dairy farmers' willingness to participate in a sustainability standard: evidence from an empirical study in Germany. International Journal on Food System Dynamics 7, 243–257.

Macpherson, J., Paul, C., Helming, K., 2020. Linking ecosystem services and the SDGs to farm-level assessment tools and models. Sustainability 12, 6617. https://doi.org/10.3390/su12166617.

Meemken, E.-M., 2020. Do smallholder farmers benefit from sustainability standards? A systematic review and meta-analysis. Global Food Security 26, 100373. https://doi.org/10.1016/j.gfs.2020.100373.

Mol, A., Oosterveer, P., 2015. Certification of markets, markets of certificates: Tracing sustainability in global agro-food value chains. Sustainability 7, 12258–12278. https://doi.org/10.3390/su70912258.

More, S.J., Marchewka, J., Hanlon, A., Balzani, A., Boyle, L., 2021. An evaluation of four private animal health and welfare standards and associated quality assurance programmes for dairy cow production. Food Policy 105, 102169. https://doi.org/10.1016/j.foodpol.2021.102169.

Osmundsen, T.C., Amundsen, V.S., Alexander, K.A., Asche, F., Bailey, J., Finstad, B., Olsen, M.S., Hernández, K., Salgado, H., 2020. The operationalisation of sustainability: Sustainable aquaculture production as defined by certification schemes. Global Environmental Change 60, 102025. https://doi.org/10.1016/j.gloenvcha.2019.102025.

Poppe, K., Koutstaal, H., 2020. Eco-schemes and private sustainability initiatives: Creating synergies. EuroChoices 19, 36–40. https://doi.org/10.1111/1746-692x.12250.

Rockström, J., Edenhofer, O., Gaertner, J., Declerck, F., 2020. Planet-proofing the global food system. Nature Food 1, 3–5. https://doi.org/10.1038/s43016-019-0010-4.

Rubio-Jovel, K., 2022. The voluntary sustainability standards and their contribution towards the achievement of the Sustainable Development Goals: A systematic review on the coffee sector. Journal of International Development 2022, 1–40. https://doi.org/10.1002/jid.3717.

Rueda, X., Garrett, R.D., Lambin, E.F., 2017. Corporate investments in supply chain sustainability: Selecting instruments in the agri-food industry. Journal of Cleaner Production 142, 2480–2492. https://doi.org/10.1016/j.jclepro.2016.11.026.

Samerwong, P., Bush, S.R., Oosterveer, P., 2018. Implications of multiple national certification standards for Thai shrimp aquaculture. Aquaculture 493, 319–327. https://doi.org/10.1016/j.aquaculture.2018.01.019.

Schiano, A.N., Harwood, W.S., Gerard, P.D., Drake, M.A., 2020. Consumer perception of the sustainability of dairy products and plant-based dairy alternatives. Journal of Dairy Science 103, 11228–11243. https://doi.org/10.3168/jds.2020-18406.

Sonderegger, G., Heinimann, A., Diogo, V., Oberlack, C., 2022. Governing spillovers of agricultural land use through voluntary sustainability standards: A coverage analysis of sustainability requirements. Earth System Governance 14, 100158. https://doi.org/10.1016/j.esg.2022.100158.

Van Calker, K.J., Berentsen, P.B.M., De Boer, I.J.M., Giesen, G.W.J., Huirne, R.B.M., 2007. Modelling worker physical health and societal sustainability at farm level: An application to conventional and organic dairy farming. Agricultural Systems 94, 205–219. https://doi.org/10.1016/j.agsy.2006.08.006.

Van der Linden, A., De Olde, E.M., Mostert, P.F., De Boer, I.J.M., 2020. A review of European models to assess the sustainability performance of livestock production systems. Agricultural Systems 182, 102842. https://doi.org/10.1016/j.agsy.2020.102842.