A systematic assessment of the revised Nutri-Score algorithm: Potentials for the implementation of front-of-package nutrition labeling across Europe

Edvina Hafner1,2 | Igor Pravst1,2,3

1 Nutrition Institute, Ljubljana, Slovenia
2 Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia
3 Faculty of Applied Sciences—VIST, Ljubljana, Slovenia

Correspondence
Igor Pravst, Nutrition Institute, Koprska ulica 98, SI-1000 Ljubljana, Slovenia.
Email: igor.pravst@nutris.org

Abstract
Front-of-package nutrition labeling (FOPNL) is an important public health tool, and the introduction of harmonized FOPNL in Europe is one of the most ambitious food labeling changes in decades. Nutri-Score (NS) has been considered for implementation across Europe. However, NS is subject to strong opposition, particularly from the food industry and some agricultural sectors (such as cheese and cured meat), but also from some nutrition scientists and public health professionals, which highlights that the system is not sufficiently aligned with food-based dietary guidelines and the latest scientific literature. These concerns were recently addressed in a revised version of NS (NS2023), aiming to overcome the limitations of its predecessor (NS2021). Our aim was to assess whether these limitations were addressed and to investigate their alignment with dietary guidelines. A systematic literature review identified 20 limitations of NS, assigned to 3 groups (Food-based, Component-based, and Other Dimensions of Food Quality). Subsequent assessment of NS employed a large representative branded food database of 19,510 pre-packed foods. Alignment with dietary guidelines was assessed based on agreement with the WHO Europe nutrient profile (WHOE). NS2023 was shown as notably stricter compared to NS2021 (7% fewer products received the higher grades A or B) and more aligned with WHOE ($\kappa_{NS2021} = .59$, $\kappa_{NS2023} = .65$). Overall, most (65%) of the limitations were addressed to some extent; these were mostly Food-based limitations, followed by Component-based, whereas the Other Dimension of Food Quality (processing, sustainability, portion sizes) remained mostly unaddressed. We can conclude that the revised NS2023 increased its potential for implementation across Europe. Our review identified all limitations, relevant or not, which were mentioned in the scientific literature. Therefore, some mentioned limitations may never be solvable in the scope of nutrient profiling. Others could be further addressed by adaptation of the visual presentation, but increased complexity of the labeling message would also reduce the potential of the FOPNL to support consumers...
in healthier food choices. Additional research is also necessary to explore the potential impact of the revised NS2023 on food reformulation and its perception among consumers.

**KEYWORDS**
food policy, food supply, front-of-package labeling, nutrient profiling, Nutri-Score

## 1 INTRODUCTION

Diet plays a crucial role in human health, and its significance as a determinant of public health has become increasingly important with changed dietary habits, a fast-paced lifestyle, and the rise of non-communicable chronic diseases (NCDs) (World Health Organization, 2020). Processed foods constitute a substantial part of the diet in developed countries, and their nutritional composition can vary considerably (Martini et al., 2021). Consequently, empowering consumers to navigate through the abundance of branded food products and support nutritionally favorable choices becomes crucial for safeguarding public health (World Health Organization, 2019).

Although detailed composition information on food products can be overwhelming for consumers (Miller & Cassady, 2015), simplified nutrition information systems in the form of front-of-package nutrition labeling (FOPNL) have been recognized as an important public health tool not only by researchers but also by policy makers (Jones et al., 2019). The European Commission therefore proposed the implementation of harmonized mandatory FOPNL by the end of 2022, but the decision was postponed due to challenges encountered in selecting the appropriate approach and profiling model (European Parliament, 2023), which was also influenced by some member states (especially Italy) and lobbying from the food industry and its supporters (Delhomme, 2021). Nutri-Score (NS) has been considered one of the main candidates for the implementation across the European Union (EU). It was created in France with the aim of providing consumers with user-friendly nutritional information to encourage healthier food choices. NS is a summary-graded FOPNL, grading products on a five-color scale marked from A (healthier) to E (less healthy) (Julia & Hercberg, 2017). The algorithm of NS was built on the United Kingdom Food Standard Agency nutrient profiling (NP) system (UK Department of Health, 2011), which was adapted by the French High Council for Public Health (Julia, 2020). NS is recognized as one of the most researched FOPNLs. Studies carried out in various settings have shown that NS is easily understood by consumers and that it can improve people's food choices and purchases (Egnell et al., 2018; Julia et al., 2021; Packer et al., 2021). Epidemiological studies have also linked a higher consumption of lower NS grade foods with an increased risk of non-communicable diseases (Donat-Vargas et al., 2021; Gomez-Donoso et al., 2021). For these reasons, several European countries, including France, Germany, and Belgium, have already adopted the voluntary use of NS (French Ministry of Health and Prevention, 2021). Proposing mandatory harmonized FOPNL across Europe would represent one of the most ambitious changes in food labeling in decades. However, the EC’s proposal faced opposition from various stakeholders (Delhomme, 2021), particularly the food industry and certain agricultural sectors (such as cheese and cured meat) (Fialon et al., 2022; Stiletto et al., 2023). Additionally, some nutrition scientists and public health experts expressed concerns about the system’s inadequate alignment with food-based dietary guidelines (FBDG), potentially leading to conflicting information for consumers (Konings et al., 2022). The proposal of NS as a potential mandatory FOPNL in Europe triggered substantial concerns and criticisms, primarily centered around the NP algorithm and its grading system. Although some criticisms were assigned as political (Julia et al., 2022), others referred to the scientific evaluation of NS alignment with FBDG (Borg et al., 2021; Konings et al., 2022).

Most studies that have examined the alignment of NS with national dietary guidelines have found that this model is generally well aligned with the guidelines but also identified opportunities for improvements (Donrea-Trecant et al., 2020; Hafner & Pravst, 2021; Szabo de Edelenyi et al., 2019). This frequently included cheeses, oils, and some types of beverages (Bourges, 2021; Hafner & Pravst, 2023). Research has also uncovered loopholes allowing certain products to mask negative attributes, achieving healthier NS grade, which is not aligned with FBDG and other FOPNLs. This has led to a growing demand for the NS to be updated to align more closely with the latest nutritional guidelines. As NS relies on a nutrient profile model that was created in 2004–2005 and adapted for FOPNL in 2015, the Scientific Committee of NS (ScC) acknowledged potential limitations that emerged since the implementation. The regular evaluations of NS are also included in its regulatory framework (French Ministry of Health and Prevention, 2021), following the example of other FOPNLs, such as Health Star Rating (mpconsulting, 2019). Therefore, to further align NS with FBDG, revision of NS was discussed between the ScC and Countries Officially Engaged in the NS (COEN), and a variety of scenarios for the update of the NS were tested (French Ministry of Health and Prevention, 2021). Following discussions, various FBDGs, and scientific literature, the ScC very recently issued two scientific reports providing a major revision of NS (NS2023) (Scientific Committee of the Nutri-Score, 2022, 2023). It should be mentioned that the discussion on the revision of NS was mostly carried out between COEN member countries (French Ministry of Health and Prevention, 2021), which are more supportive of the implementation of NS. The limited inclusion of individuals from other countries might have resulted in overlooking profiling issues raised by those less supportive of the system. Although some criticisms of NS2021 were commonly mentioned, others may be hidden in the
extensive scientific literature on NS, which has, to our knowledge, not been systematically reviewed before. Although there are some studies that have analyzed the revised NS model (Øvrebø et al., 2023; Sarda et al., 2024), we were not able to find a study that evaluated improvements in the model through a direct comparison between the original NS2021 and the revised NS2023 using real-world food supply data and subsequently assessing changes in alignment with dietary guidelines.

The aim of our study was therefore to systematically review all reported limitations of the NS2021 system in the scientific literature (not limited to COEN countries) and to assess whether these limitations were addressed in the 2023 revision of the NS based on a large, branded food database, comprising 19,510 pre-packed foods. Furthermore, we investigated the alignment of the NS2023 profiling algorithm with dietary recommendations, as reflected by the WHO Europe nutrient profile model (WHOE) (Ministry of Health Republic of Slovenia, 2016). It should be noted that in the absence of an EU-representative branded food database, this study was carried out with a dataset representative of the food supply in Slovenia, an EU member state, with a major proportion of imported branded foods from other European countries. Slovenia is also not one of the COEN countries that were included in the revision of NS, making the study results also relevant to regions where NS is not yet implemented.

2 METHODS

2.1 A systematic scientific literature review of NS limitations

A systematic scientific literature search was carried out following the Preferred Reporting Items or Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009), presented in Figure 1.

The aim of the systematic review was to identify papers that in any way address the limitations of the NS algorithm and/or its misalignments with dietary recommendations and the scientific literature. Consequently, we excluded papers that reported the limitations of NS in visual appearance, consumer perception, attention, and similar. A Boolean search was conducted using the selected terms (“nutriscore” OR “nutri-score” OR (“nutri” AND “score”)) on June 26, 2023, with no specific start date. The search was run for peer-reviewed publications (titles, abstracts, and keywords) in the Web of Science and Scopus databases. Studies that met any of the following criteria were excluded from the review: (1) The full text was not available; (2) the study analyzed the impact of NS on consumers; (3) the study analyzed the effect of NS on health outcomes; (4) the study used NS only as a tool to evaluate the healthiness of food products; (5) the full paper did not mention any limitations of the NS system. After this screening, 30 publications were included in the review (Table S1). The papers were brought together in clusters, based on the reported types of limitations. We identified three clusters of reported limitations: Food-based (connected to alignment with FBDG), Component-based, or Other Dimensions of Food Quality. As each limitation was placed in the wider context of nutrition, we found connections among the revealed limitations, which were presented in a network graph.

2.2 Dataset and food categorization

A large and representative branded food database is required for meaningful assessment of the NP algorithms. The EU currently lacks a comprehensive branded food database, and only a handful of EU countries gather such data, primarily focusing on specific food categories. Slovenia is one of the exceptions; extensive systematic monitoring of the branded foods in the food supply was introduced here in 2011, with Composition and Labelling Information System database managed by the Nutrition Institute (Ljubljana, Slovenia). Database includes data from cross-sectional studies done every couple of years in major food retailers in Slovenia to make monitoring of the food supply up-to-date (Pravst et al., 2022). Dataset for each year represents offers at the time of sampling. For this study, we employed the latest available representative dataset from 2020. The detailed methodology on data collection is described elsewhere (Pravst et al., 2022); in summary, the methodology includes the collection of photographs of food labels, which were available in all major grocery stores (without being limited to specific food categories). Photographs were then used for the extraction of data on the nutritional composition, ingredients, descriptive name, and so on. Each product in the database is identified by the Global Trade Item Number (GTIN), which prevents duplication of products in the database. The complete 2020 dataset included 28,028 pre-packed foods and beverages, which were classified using the international food categorization of the Global Food Monitoring group (Dunford et al., 2012).

After the data collection, the dataset went through data cleaning (Hafner & Pravst, 2021). In brief, data cleaning included the exclusion of products that could not be profiled by NS. Most products were excluded because they are not covered by NS systems (n = 6966). This mostly included any type of alcoholic beverages, teas, coffee, and so on. We also excluded products with missing mandatory information (n = 881) and products with a preparation process that includes the use of other food ingredients (n = 546). To minimize errors in the recorded nutritional composition, we also excluded products for which the calculated energy notably deviated (± 20%) from the labeled energy value (n = 125). Energy values were calculated in line with the methodology provided in Regulation (EU) No. 1169/2011 (European Commission, 2011). After the exclusion process, the final dataset used for NP in this study consisted of 19,510 pre-packed products, categorized into 13 main categories and 55 subcategories.

As we mentioned, Slovenia is part of the European open market, and a majority of food products are imported from other European countries. To identify imported products in our database, we conducted analyses based on GTIN codes. The presence of the Slovenian country code (indicating that the company is registered with the GS1 association in Slovenia, which assigns GTIN codes for Slovenia) was identified in 31% of the cases. Based on the result, we can thus infer that 69% of the products in our database is imported. This result closely aligns
FIGURE 1 Preferred Reporting Items or Systematic Reviews and Meta-Analyses (PRISMA) flow chart of the systematic review of Nutri-Score (NS) limitations.

with the information provided by the Statistical Office of Republic of Slovenia, which in the 2021 report for processed food states that 72% of such food was imported (Statistical Office of Republic of Slovenia, 2022).

2.3 Nutrient profiling with NS and WHOE

Each product from the final dataset \( n = 19,510 \) underwent NP three times: using the original NS algorithm (NS2021), the revised NS algorithm (NS2023), and the 2015 edition of the WHOE algorithm (World Health Organization, 2015), which was also nationally implemented (Ministry of Health, Republic of Slovenia, 2016). Profiling with NS2021 was carried out according to the Scientific and Technical instructions of the French Public Health Agency (Hafner & Pravst, 2023; Santé publique France, 2021). Profiling with NS2023 adhered to the latest revision of the algorithm, which is described in two reports (update on solid foods [Scientific Committee of the Nutri-Score, 2022], and update on beverages [Scientific Committee of the Nutri-Score, 2023]). In brief, NS is determined based on the final sum of positive and negative points that represents positive/negative attributes of nutritional composition (per 100 g/mL) of a product. Positive points represent negative attributes (energy value, sugars, saturated fats, and salt), whereas negative points represent positive attributes (% fruit vegetables and legumes, protein, and dietary fiber). The sum of all positive and negative points gives the final score of a product, which determines a grade from A (healthier) to E (less healthy). The main differences that were made with revised NS2023 include changed rating scales and thresholds for some nutrients (salt, sugar, fiber, and protein), changed definition of % fruit, vegetables, and legumes (%FVL) (excluding % of nuts and olive, rapeseed, and walnut oil), added special negative attributes (red meat and non-nutritive sweeteners [NNS]), changed profiling category for nuts and dairy beverages, and changed thresholds for the final grade. More detailed description of differences between original and revised NS algorithms is available in reports from ScC (Scientific Committee of the Nutri-Score, 2022, 2023).

The NP of the products included some information beyond mandatory food composition labeling; composition data was therefore supplemented according to previously reported methodology (Hafner & Pravst, 2021, 2023). The data was supplemented by dietary fiber content, content of red meat, and the percentage of FVL (%FVL). For products without labeled dietary fiber content, the content was calculated for all relevant categories based on their energy value and other nutrient content (European Commission, 2011). An assessment of the %FVL was estimated from ingredient lists, employing regulation requirements for specific food groups (Republic of Slovenia, 2004, 2013). This process included 11,352 products, with the possibility of having sufficient %FVL, relevant for NS (>40%FVL). In the end, %FVL was determined for 2820 products, with 1571 of them containing 100% FVL (such as juices, dried and frozen fruit/vegetables,
and dried legumes). The content of red meat was determined from ingredient lists, following the protocols of each profiling method (Ministry of Health, Republic of Slovenia, 2016; Scientific Committee of the Nutri-Score, 2022).

The profiling with WHOE was used as a proxy of comparing NS to FBDG based on several factors: (a) dietary guidelines in the EU can vary between different member states; therefore, opting for a neutral system that is used in several member states was deemed the most appropriate approach. We also should note that in Slovenia, FBDGs are not defined to a level that would allow meaningful comparison with NS. On the other hand, WHOE is not country-specific and is therefore more relevant for comparison with NS, which is also used in several different countries. The WHOE profile was also developed based on available NP models in Europe, and “nutrient thresholds were introduced following consultation with countries and to be in line with WHO nutrition guidelines” (World Health Organization, 2015). (b) Moreover, the WHOE is an NP model organized by food categories, in contrast to NS, which follows an across-the-board approach. The ample variety of food categories makes WHOE closely aligned with FBDG. (c) Direct quantitative comparison is possible between NS and another profiling system, which would not be the case for comparison with FBDG. (d) WHOE was developed to restrict food marketing to children, which are one of the most vulnerable population groups, which should also be included when developing the NP model and FOPNL (Storchsdieck genannt Bonsmann et al., 2020). This also positions WHOE as one of the more rigorous NP models when compared to other models (Pivk Genannt Bonsmann et al., 2020). The WHOE was also implemented into regulation in Slovenia, where it is used to limit the marketing of food products (Ministry of Health, Republic of Slovenia, 2016).

However, certain limitations should be acknowledged in making such comparisons. It is important to note that despite WHOE serving as a proxy, this system has its own unique constraints and may not consistently align with the latest FBDG, particularly in specific areas (e.g., red meat and whole grains). These issues are further addressed in the discussion section.

2.4 Data analyses

The data analyses were conducted using R: A Language and Environment for Statistical Computing 2020 (R Core Team) and Microsoft Excel 2019 (Microsoft).

First, we examined the differences in the distribution of the different grades between NS2021 and NS2023 at the (sub)category level. Next, we determined the proportion of products where the final NS grade of a product changed between NS2021 and the revised NS2023. In each of the 55 subcategories, we further divided the products into narrower groups based on 3 different aspects: the descriptive name of the food, ingredients, and critical nutrients. Based on this, we qualitatively identified more specific products that were most exposed to changes in the NS algorithm.

To evaluate changes in alignment between NS and WHOE, we calculated Cohen’s kappa and the percentage of agreement between NS2021/NS2023 and WHOE. The comparison was conducted by dividing products into “healthier” and “less healthy” options based on the following criteria, which determined “healthier” products: grade A or B for NS, and permitted for marketing to children for WHOE (Hafner & Pravst, 2021; Poon et al., 2018).

3 RESULTS

3.1 Reported limitations of NS2021 algorithm

Our literature review captured 30 papers: 24 original research (Angelino et al., 2023; Braesco et al., 2022; Bryngelsonn et al., 2022; Cutroneo et al., 2022; de las Heras-Delgado et al., 2023; Dickie et al., 2022; Ebner et al., 2022; Fedde et al., 2022; Ferreiro et al., 2021; Fialon et al., 2022; Hafner & Pravst, 2021, 2023; Huybers & Rodenburg, 2023; Katsouri et al., 2021; Kissock et al., 2022; Konings et al., 2022; Panczyk et al., 2023; Pitt et al., 2023; Pointcke & Pawelzik, 2022; Rodriguez-Martín et al., 2023; Septia Irawan et al., 2022; Stiletto & Trestini, 2022; Borg et al., 2021; Valenzuela et al., 2022), 3 communication/correspondence (Carruba et al., 2022; Drewnowski et al., 2021; Hau & Lange, 2023), and 3 review papers (Peonides et al., 2022; van der Bend et al., 2022; Włodarek & Dobrowolski, 2022), originating from 19 countries (including 12 EU member states). The list of reported limitations indicates that (n = 20), their connections, and the numbers of mentions are presented in Figure 2. Overall, we observed that most reported limitations were Food-based, followed by Other Dimensions of Food Quality and lastly by Component-based limitations.

The most commonly reported limitation (n = 15) was part of Other Dimensions of Food Quality, which was the lack of consideration of ultra-processed foods (UPFs). UPFs can sometimes achieve a better NS grade than less processed foods (Angelino et al., 2023; Ebner et al., 2022; Fedde et al., 2022; Hau & Lange, 2023). This was also commonly reported with reference to Food-based limitations such as oils (n = 10) (Hau & Lange, 2023; Septia Irawan et al., 2022) and cheeses (n = 9) (Ebner et al., 2022; Fialon et al., 2022), which can be less processed than UPFs but often get lower grades, which may indicate a noncompliance of NS with FBDG. Similar arguments were mentioned for beverages (n = 8), where fruit juices usually get a lower NS grade than ultra-processed drinks with added NNS (Włodarek & Dobrowolski, 2022).

Another frequently raised limitation within the Other Dimensions of Food Quality cluster pertained to the suitability of NP using a uniform/fixed quantity of food across all categories (per 100 g/mL). This was often mentioned considering Food-based limitations, especially for oils, cheeses, and nuts, which are energy-dense foods typically eaten in smaller portions (Braesco et al., 2022; Fialon et al., 2022). Additionally, this issue was raised in relation to ready meals, which are often eaten in portions much larger than 100 g and can therefore achieve a better NS grade (Fialon et al., 2022; Hau & Lange, 2023). Oils and cheeses are also often advocated in the light of traditional foods (Other Dimensions of Food Quality), for example, as part of a Mediterranean diet, which, based on the studies, could be possibly impacted by the NS algorithm (Fialon et al., 2022; Peonides et al., 2022).
Most common Component-based limitation was masking negative components (mostly high salt and sugar content) by the addition of positive components, such as dietary fiber, protein, and %FVL. This limitation was closely connected to many Food-based and Other Dimensions of Food Quality limitations. Studies particularly highlighted masking high sugar content in breakfast cereals with fiber (van der Bend et al., 2022), high sugar content in yogurts by adding protein (Włodarek & Dobrowolski, 2022), and high salt content in plant-based alternatives with % FVL and dietary fiber (Huybers & Roodenburg, 2023). Authors also suggested that both masking and reformulation are much easier for UPFs than for staple foods or traditional specialty guaranteed products (Fialon et al., 2022), and that NS is thus favoring UPFs. Another Component-based limitation that the review highlighted was that the NS algorithm does not sufficiently account for the content of whole grain ingredients, which are commonly recommended in dietary guidelines (Blomhoff et al., 2023). For example, unprocessed whole grain breakfast cereals may get the same NS grade as UPF breakfast cereals with added fiber and protein. Although these two products may have a similar nutritional composition, the processing of cereals can make the starch matrix unstructured, which can result in a higher glycemic index of products in comparison to whole grains. Furthermore, UPFs may not have the same micronutrient value as whole grains. Therefore, same grading for such different products may be inappropriate (Ebner et al., 2022).

A lack of inclusion of micronutrients and other components into NS is another Component-based limitation mentioned in various contexts connected to Food-based limitations and UPFs. Examples refer to the importance of calcium in dairy products (Huybers & Roodenburg, 2023), phenolic components, fatty acid composition in olive oil, vitamin content of juices, and omega-3 fatty acids in fish (Włodarek & Dobrowolski, 2022). Some authors also argue that plant-based alternatives, which are also mostly UPFs, may not have a comparable content of micronutrients to meat and dairy products, and that the bioavailability of micronutrients should also be considered (Huybers & Roodenburg, 2023). One of the mentioned Component-based limitations is also that the NS algorithm does not consider the addition of NNS; this was mostly mentioned for beverages, yogurts, and dairy drinks, where products sweetened with NNS can get the same or an even better grade than less processed (plain) products (Borg et al., 2021; Fedde et al., 2022). All the other limitations were mentioned less than three times and were mostly Food-based (nuts, yogurts, cereals, red meat, sauces, and fish) and connected with the previously discussed Component-based or Other Dimensions of Food Quality limitations. Among other limitations, some studies argued that NS does not take into account the carbon footprint and organic production/pesticide residues of the food (Panczyk et al., 2023).

### 3.2 Changes in the distribution of Nutri-Score grades

Next, we accessed the distribution of NS2021 and NS2023 grades across our dataset. The distribution for the main categories is displayed in Figure 3, whereas the results for subcategories are presented in Table S2. Overall, NS2023 was a notably stricter profiling model than NS2021. For N2021, the most common grade for products was D.
FIGURE 3  Distribution of the original (NS2021) and updated (NS2023) Nutri-Score grades across the main categories, based on the 2020 Slovenian food supply (n = 19,510). Exact values and subcategory results are available in Table S2. Letters from A to E and the corresponding colors represent five Nutri-Score grades.

(28%), whereas for NS2023, this shifted to E (32%). Although the proportion of products with grade C remained relatively unchanged, there was a notable reduction in the “healthier” grades A and B (NS2021: 30%, NS2023: 23%).

For beverages, we overall observed improved grading. NS2023 graded significantly fewer products at E (32%) than NS2021 (41%). NS2023 graded more beverages at the middle grade, C (32%), which was less prevalent with NS2021 (17%). With NS2023, grade C became very diverse, including different types of beverages, from sugar-sweetened drinks to 100% juices and drinks with NNS. NS2023 graded fewer beverages at B (6%) compared to NS2021 (11%). For bread and bakery products, we observed that NS2023 graded more products at D and E than NS2021; this was most notable for the subcategories: biscuits and cakes, muffins, and pastry. In the bread subcategory, the predominant grade for NS2021 was B (46%), whereas with NS2023, it has shifted to C (61%). Most breads were downgraded with NS2023, making grade A accessible for only 9% of the products, compared to 28% with NS2021. For cereal and cereal products, NS2023 was stricter in grading products A or B. This mostly impacted breakfast cereals, where offsetting high sugar levels with positive attributes is still possible but much more challenging with NS2023. For confectionery, NS2023 exhibited stricter grading, with 81% of products receiving grade E, compared to 56% in NS2021. Convenience foods showed the largest disparity in the proportion of healthier products (A or B). NS2021 classified nearly half of the products (47%) as healthy, whereas NS2023 limited such grading to only 26%. Downgrading was evident across all the subcategories, especially for pizza. For dairy and alternatives, we observed that NS2021 classified 35% of products as healthier options, whereas NS2023 reduced this to 15%. This was particularly notable for milk and milk drinks and milk alternatives, which were in different profiling categories with NS2023. This resulted in overall downgrading, with B being the highest grade in NS2023. In contrast, NS2021 graded 35% of milk and milk drinks and 18% of milk alternatives at A. Similar downgrading was observed for all yogurt subcategories: plain yogurt, flavored yogurt, and yogurt alternatives. Edible oils and emulsions were the most impacted by the NS2023 changes. NS2021 graded products mostly at C (43%), followed by D (39%) and E (18%), whereas NS2023 allows some products to get the more favorable grade B. Grade C (52%) remained the most common, followed by B (22%) and E (21%), whereas only a few products were graded D (5%). For eggs, NS2023 showed improved grading, with almost all products being graded at A. In the category of fish and fish products, we observed a more balanced distribution of grades when using NS2023. There was a notable increase in the proportion of A-graded products, particularly for unprocessed chilled fish (NS2021: 16%, NS2023: 22%). However, there was an increase in the proportion of products graded E (NS2021: 2%, NS2023: 12%), mostly due to the downgrading of products in the processed chilled fish products subcategory. NS2023 was stricter than NS2021 for fruit and vegetables. Changes were noticeable in the subcategories of dried fruit, nuts and fruit mixes, and jam and spreads, where NS2023 graded more products at D (43%, 78%, and 75%, respectively), and even some at grade E (14% of nuts and fruit mixes), which was not found in NS2021. In meat, meat products, and alternatives, there was a higher proportion of products graded E with NS2023 (55%) than with NS2021 (44%), primarily due to the stricter profiling in the subcategory of processed meat and meat products. Sauces and spreads were downgraded with NS2023,
TABLE 1  Agreement (Cohen’s kappa and % of agreement) of the WHO Europe profile with the previous (NS2021) and updated (NS2023) Nutri-Score based on the 2020 Slovenian food supply (n = 19,510).

<table>
<thead>
<tr>
<th>Category</th>
<th>Cohen’s kappa</th>
<th>% of agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NS2021</td>
<td>NS2023</td>
</tr>
<tr>
<td>All products (n = 19,510)</td>
<td>0.59</td>
<td>0.65</td>
</tr>
<tr>
<td>Beverages (n = 1881)</td>
<td>0.39</td>
<td>0.52</td>
</tr>
<tr>
<td>Bread and bakery products (n = 2304)</td>
<td>0.73</td>
<td>0.57</td>
</tr>
<tr>
<td>Cereal and cereal products (n = 2008)</td>
<td>0.78</td>
<td>0.80</td>
</tr>
<tr>
<td>Confectionery (n = 2283)</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Convenience foods (n = 750)</td>
<td>0.53</td>
<td>0.46</td>
</tr>
<tr>
<td>Dairy and alternatives (n = 3203)</td>
<td>0.44</td>
<td>0.46</td>
</tr>
<tr>
<td>Edible oils and emulsions (n = 587)</td>
<td>na</td>
<td>0.18</td>
</tr>
<tr>
<td>Eggs (n = 69)</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Fish and fish products (n = 577)</td>
<td>0.63</td>
<td>0.63</td>
</tr>
<tr>
<td>Fruit and vegetables (n = 2147)</td>
<td>0.50</td>
<td>0.68</td>
</tr>
<tr>
<td>Meat, meat products, and alternatives (n = 1890)</td>
<td>0.53</td>
<td>0.47</td>
</tr>
<tr>
<td>Sauces and spreads (n = 1207)</td>
<td>0.23</td>
<td>0.35</td>
</tr>
<tr>
<td>Snack foods (n = 604)</td>
<td>0.40</td>
<td>0.63</td>
</tr>
</tbody>
</table>

having a higher proportion of products graded D and E. This was the case for all subcategories except nut spreads, which were graded higher with NS2023 than with NS2021, due to changed profiling for nuts and their products. Finally, snack foods were also downgraded with NS2023, with grades D and E representing 89% of all the products, whereas for NS2021, this was 74%.

In the next step, we evaluated the proportion and types of products that will change their NS grade with the implementation of NS2023. We observed that over one third of the products (34%) changed their grade from NS2021 to NS2023; overall, 28% worsened their grade and 6% improved it. Figure 4 shows the proportions of products that changed their NS from NS2021 to NS2023 in each main category. Where possible, we also identified the most common foods within the category, which changed their NS for better or worse. Categories with a notable proportion of products with an improved NS grade were edible oils and emulsions, eggs, beverages, and fish and fish products. In other categories, approximately one third of the products experienced changes in their NS, mostly indicating a downgrade. The category least affected by these changes was meat, meat products, and alternatives, where 19% of products saw a shift with NS2023, with 3% improving and 16% worsening.

3.3 | Agreement with dietary recommendations

The harmonized WHOE NP model was used as an indicator of whether specific food products could be recommended for regular consumption in line with dietary recommendations. Although NS2021 showed moderate agreement with the WHOE profile ($\kappa_{NS2021} = .59$) (Table 1), which is in line with previous reports (Hafner & Pravst, 2021), NS2023 showed improved, strong agreement ($\kappa_{NS2023} = .65$) with the WHOE profile. In most categories (7 out of 13), NS2023 showed higher agreement than NS2021. Slightly lower agreement was shown for bread and bakery products ($\kappa_{NS2021} = .73$, $\kappa_{NS2023} = .57$), convenience foods ($\kappa_{NS2021} = .53$, $\kappa_{NS2023} = .46$), and meat, meat products, and alternatives ($\kappa_{NS2021} = .53$, $\kappa_{NS2023} = .47$). We should mention that these are categories, where NS2023 was stricter compared to WHOE, and where WHOE somehow lacks alignment with latest FBDG (importance of fiber in breads and limiting consumption of red meat). Agreement was particularly improved for edible oils and emulsions, where no product passed the “healthier” threshold with NS2021 (% of agreement with WHOE 26%); this changed with NS2023, where 22% of products were identified as “healthier” (% of agreement with WHOE 48%). As has been shown before (Hafner & Pravst, 2021), the WHOE profile was demonstrated to be stricter than NS2021, but this has changed with NS2023, particularly for dairy and alternatives. In our dataset, WHOE classified 28% of products as healthier, compared to 30% for NS2021 and 23% for NS2023 (Table S2).

3.4 | Did the NS revision address the previously reported limitations?

Our results indicate that the last revision of the NS addressed at least some extent of 65% of the limitations identified for NS2021 (13 out of 20) (Table S3). The revision particularly focused on Food-based limitations (11 out of 12), except for whole grains. It also tackled most of the Component-based limitations (2 out of 3), including the masking of negative components and the penalization of NNS, whereas the inclusion of micronutrients in the profiling was not part of the updated algorithm. The update did not cover any of the limitations falling under the category of Other Dimensions of Food Quality (0 out of 5). The extent to
which the updated algorithm addressed each specific limitation of its predecessor is described as follows.

Food-based limitations cluster, which represents a lack of alignment with FBDG, was most extensively addressed. Based on our results, we can say that oils were addressed by the improved distribution of NS grades across the category, which also resulted in higher grades for olive oil. Cheeses' limitations were partially addressed by a more even distribution of NS grades across the category, with a slightly higher proportion of C-graded and a higher proportion of E-graded cheeses. This allows greater differentiation between products. However, critiques suggesting that cheeses should receive higher grades (with some attaining healthier grades A or B) remains unaddressed, resulting in
most cheeses still receiving an unfavorable grade D, but the stricter algorithm for cheeses did not address some criticism with regard to low grading compared to other dairy products. The limitations about beverages were also addressed, particularly in connection with added NNS (Scientific Committee of the Nutri-Score, 2023). Based on our results, beverages with NNS lowered their grade and achieved a maximum grade C. Based on that, we also observed that NS2023 mostly differentiates between naturally low-calorie drinks, such as flavored waters with a small amount of fruit juice (graded B or C) and drinks with added NNS (graded C or D).

The possibility of masking negative components (Component-based limitation) was addressed very comprehensively in the latest revision of NS. In NS2023, sugar and salt content are more penalized, whereas the conditions for obtaining fiber and protein points are higher. This was particularly notable for some Food-based limitations; for example, breakfast cereals with high sugar content can no longer easily mask their nutritional composition to get a grade A. While still possible, masking of high sugar content in these products is much more challenging in comparison with NS2021. Such products were downgraded by one or even two grades. Higher penalization of sugar and salt resulted in an overall stricter algorithm, which also addressed limitations for ready meals, yogurts, sauces, and plant-based alternatives, which notably downgraded. Added NNS was addressed by adding it as a negative attribute in beverages and dairy drinks, which consequently downgraded by one grade; however, it was not addressed in solid yogurts, where products with added NNS can still achieve a “healthier” grade (A or B). NS2023 also addressed the profiling limitations of plain nuts, which mostly improved their grade; this also resulted in notably improved agreement with the WHOE profile (\( r_{NS2021} = .46 \) to \( r_{NS2023} = .72 \)). To make the algorithm more aligned to the latest nutritional guidelines, red meat was penalized, which gave NS2023 a stricter modeling profile than NS2021. However, it did not notably change the agreement with the WHOE profile, which does not have specific rules referring to red meat content. For fish, we observed better discriminatory ability, especially for less processed products, which improved their grades from B to A.

NS2023 did not directly cover the limitations referring to the content of micronutrients and whole grains. With stricter profiling and an adaptation for fiber scoring, NS2023 demonstrates improved capability in distinguishing between breads compared to NS2021, based on fiber content. However, we should note that the algorithm does not explicitly account for whole grains. Consequently, certain breads containing added seeds or protein may receive the same grade as whole-grain breads. Given the absence of a legal definition for whole grains in the EU (European Commission, 2023), this aspect should be further assessed and falls beyond the scope of this study. The limitations of Other Dimensions of Food Quality were also not covered by NS2023. However, although limitation of UPFs has not been directly addressed by modified NS2023, it might be addressed indirectly, which could be further researched with classification systems that address level of processing (like NOVA), which were not part of our study.

4 | DISCUSSION

The present study has focused on all the reported limitations (relevant or not) of the NS NP algorithm and how they were tackled in the recent revision (NS2023). Based on the NP results of a representative branded foods dataset (CLAS 2020), we observed that, overall, NS2023 is a notably stricter profiling system than NS2021, and that the latest revision of NS has addressed the majority of the limitations identified in the literature review. However, it is noteworthy that the extent of these improvements varied markedly across the different limitations, also based on their relevance.

4.1 | Food-based limitations

Food-based limitations were commonly related to the fact that most countries have implemented FBDG and therefore addressed the lack of alignment between NS and FBDG. It was also one of the main reasons for NS revision, as improving alignment with FBDG around EU member states is essential for further implementation on the market (World Health Organization, 2019). However, although both FBDG and NP are based on strong scientific evidence and can have synergic effect on consumers eating habits (Borg et al., 2021), they are still methodologically very different, making some discrepancies between the two approaches unavoidable (Dickie et al., 2022). On the other hand, both approaches aim to inform and educate consumers on healthy diets. Therefore, we should aim to resolve major issues to empower consumers rather than potentially confuse them with conflicting information. Such confusion could undermine the credibility of both FOPNL and FBDG (van der Bend et al., 2022). One possible approach already mentioned is a hybrid method, where nutrient density is combined with selected food groups (Drewnowski et al., 2021), as in the food category–based WHOE profiling model (World Health Organization, 2015) that was used in our study as a proxy for FBDG. We observed that NS2023 showed a better agreement with WHOE than NS2021. The agreement rose from moderate to strong, showing that NS also indirectly resolved many of the Food-based limitations (11 out of 12). Before NS revision, there were several studies that compared NS2021 with FBDG, which highlighted that NS has overall good alignment with FBDG, especially when it comes to fruit, vegetables, pulses, and unsalted nuts. It was less aligned when it came to meat, refined grains, and sweetened dairy products (Borg et al., 2021). In the latter areas, NS2023 became stricter, addressing exposed limitations of its predecessor. In some areas, such as for red meat, NS2023 became more aligned with latest dietary guidelines (Blomhoff et al., 2023; World Health Organization, 2023) than WHOE. However, as reported elsewhere (Ovrebe et al., 2023), some minor inconsistencies still remain with NS2023. Even though NS resolved many of the Food-based limitations, it remains an NP with across the board approach, which limits its alignment to FBDG to some extent. Although transitioning NS to a food category approach could address some additional limitations, such changes were not feasible with this update. The ScC
was mandated that in the updated version of NS, they shall not modify the core principles of NS, such as food categories, the calculations per portion, or the inclusion of nutrients that are not part of the mandatory nutritional declaration such as vitamin and minerals (The Luxembourg Government, 2021). For Food-based limitations, this means that some limitations cannot be addressed directly. This includes whole grains that have been addressed in some NP models (e.g., whole grains were addressed within the Keyhole FOPNL system [Pitt et al., 2023]) but are not covered by NS2023. A major challenge for considering the content of whole grains in the profiling model is the lack of the legal definition in the EU (European Commission, 2023). Furthermore, the content of whole grain ingredients is not part of mandatory food labeling, and they cannot be measured by laboratory techniques.

### 4.2 Component-based limitations

Although NS2023 improved discrimination between healthier and less healthy products in most cases, it did not directly address some limitations (e.g., added sugar and whole grains). This leaves room for algorithm manipulation (i.e. with adding protein, dietary fiber), as reported previously for certain products like breakfast cereals (van der Bend et al., 2022). Although these issues could be regarded as Food-based limitations, they could also be considered Component-based limitations. Connection between these two clusters of limitations was especially evident in our literature review. Component-based approach for revision might be more feasible because NS NP is also based on specific constituents—particularly when these are part of mandatory food labeling. With NS2023, the majority of Component-based limitations found by our review were addressed with stricter profiling, which included higher penalization of sugar, salt, and additional penalization of NNS. This adjustment does not completely prevent manipulation of the algorithm, but based on our current results, masking negative attributes has become much more challenging with NS2023. It is important to note that such manipulations are more prevalent in products where multiple nutrients collectively influence the final NS grade (breakfast cereals, convenience foods, and some dairy products). In beverages, the primary factor influencing the NS grade is often a single-key nutrient, mainly sugar, limiting options for reformulation (Zupanić et al., 2020). Conversely, breakfast cereals exhibit greater diversity (Angelino et al., 2023; Devi et al., 2014), encompassing various key constituents, including sugar, saturated fats, fiber, and protein, contributing to the final grade. Although reformulating breakfast cereals by increasing fiber or protein may enhance their nutritional profile and NS grade, it does not address the core issue of excessive added sugar (Prada et al., 2021). We must mention that, similarly to whole grains, added sugar is also not part of mandatory food labeling (European Commission, 2011); therefore, it would be very challenging to include it into NP. Such limitations would be even more relevant if mandatory use of FOPNLs is implemented. Effect of NS on food reformulation and environment should be investigated more in detail when revised NS2023 will be fully implemented on the market. One of the major knowledge gaps is the lack of in-depth studies about how such reformulation practices (possibility for algorithm manipulations) are used in connection to marketing and voluntary use of FOPNL (Storcksdieck genannt Bonsmann et al., 2020). Will reformulation follow the voluntary FOPNL, or will it be the opposite, and FOPNL will be used as a marketing tool, after reformulation? Another aspect that was highlighted before is that food producers showed a preference for Guideline Daily Amounts (GDA) labeling over multiple traffic light systems as the GDA uses portion sizes that (when adjusted) can cover up higher amounts of unfavorable characteristics (Van Camp et al., 2010). Mandatory harmonized FOPNL, for example, could limit such uses of FOPNL and more efficiently encourage producers to reformulate their products.

One of the Component-based limitations that have not been addressed in the latest NS revision is the inclusion of micronutrients. Even though the European Food Safety Authority encourages the inclusion of some micronutrients (e.g., potassium [European Food Safety Authority, 2022]) in NP, such an approach seems challenging for both the industry and the authorities, as in the EU the content of micronutrients is also not declared mandatory, with the exception of sodium/salt, and micronutrients in fortified foods (European Commission, 2011). Therefore, these system limitations should be approached with caution; a much broader change of regulatory framework would be needed before they can be efficiently incorporated into a NP system intended not only for research purposes, but also for market implementation.

### 4.3 Other Dimensions of Food Quality and constraints of nutrient profiling

Within the scope of Other Dimensions of Food Quality, we observed a wide spectrum of limitations that cover important parts of diet but might be irrelevant from the perspective of NP, which pertains only to a segment of diet (Scarborough & Rayner, 2014). Such limitations presented around one third of the reported limitations, but none of them were directly addressed. The main reason for this is that in NP, such dimensions are often out of the scope (Galan et al., 2021). A commonly expressed limitation was the level of processing, which is not included in the algorithm. In recent years, UPFs have been extensively investigated, and some studies have suggested that increased intake of UPFs can be linked to increased risk of NCDs (Food and Agriculture Organization of United Nations, 2019; Lane et al., 2021; Rauber et al., 2018). In this regard, UPFs have also become a part of many FBDGs that discourage their excessive consumption (Koios et al., 2022). Even though the main challenge of UPFs is commonly unfavorable nutritional composition, some studies suggest other negative impacts (Bonaccio et al., 2022). Recent study also revealed that associations between UPFs and unfavorable health outcomes might vary based on food category (Cordova et al., 2023). Given that the consumption of UPFs is increasing and constitutes a large part of the modern diet (Juul & Hemmingsson, 2015), such concerns should be under scrutiny with regard to food labeling. The developers of NS are currently working on alternatives to address such limitations, exploring...
future revisions such as incorporating a black border in the graphi-
cal presentation for all UPFs, which still needs further research (Srour et al., 2023). However, although UPFs are becoming a hot research
topic, this area is still under scrutiny by many scientists, referring to
oversimplified and/or lacking definition and questionable health
impact of UPFs (Farde & Rock, 2019). A NOVA system (Monteiro et al.,
2019) is a commonly used classification system with quite some con-
troversies (Petrus et al., 2021). Various other systems for categorizing
UPFs are also employed, resulting in a lack of harmonization (Martinez-
Perez et al., 2021). Therefore, the inclusion of UPFs in any NP algorithm
is questionable and challenging. We should also mention that the NS
committee recognized the importance of these different dimensions of
food quality but highlighted that none is exclusive and able to summa-
rize on its own the healthiness of a product (Galan et al., 2021). They
proposed empowering consumers with comprehensive information on
various dimensions of food quality by parallelly presenting different
food labeling schemes, including organic food label (Hercberg et al.,
2021; Srour et al., 2023). Such approach could address multiple limi-
tations, including pesticide residues identified in our review (Panczyk
et al., 2023; Wlodarek & Dobrowski, 2022). It should be noted that
the EU has already implemented a harmonized organic label (European
Commission, 2018), which could be easily presented side by side with
NS. However, this might lead to information overload and increased
complexity, potentially undermining the goal of easy comprehension
(Storcksdieck genannt Bonsmann et al., 2020). Studies highlighted
challenges of information overload on food labels, which are very re-
levant also for FOPNL (Mejean et al., 2013). Shifting attention from one
label to another could affect their effectiveness (Drescher et al., 2014).

Another limitation that falls in the cluster of Other Dimensions
of Food Quality is sustainability, mostly in connection to carbon footprint.
Even though labeling of sustainability (e.g., Eco-Score) (De Bauw et al.,
2021) has become very relevant in recent years, it still lacks harmoniza-
tion and validation (Kaczorowska et al., 2019). As NS is about healthier
food choices, inclusion of sustainability in the algorithm would be com-
plicated and could result in the overall simplification of the final grade,
which would reduce the value of the labels to consumers.

Critics also argue that NS penalizes traditional foods, posing chal-
enges for companies preserving traditional diets due to the difficulty
of reformulation, based on regulated recipes (Valenzuela et al., 2022).
Although the NS2023 algorithm did not specifically address this cri-
icism, the revision improved the profiling of some traditional foods,
which represent an important part of a healthy diet, such as olive oil and
cheeses with a lower salt content (Scientific Committee of the Nutri-
Score, 2022). However, NS2023 is still built on nutritional composition,
whereas some traditional foods are indeed high in salt, saturated fats,
and/or sugar and should be therefore consumed in moderation, even if
they are traditional (Trichopoulou et al., 2007). Therefore, such limited-
ations could be treated as irrelevant, as they cannot be covered in the
scope of NP and are often driven by lobbies (Julia & Hercberg, 2018).
We should also mention that NS2023 is still calculated per 100 g or
mL (and not by portion), which was also one of the reported limited-
tions. Therefore, many energy-dense foods eaten in smaller portions
get lower grades (Carruba et al., 2022; Katsouri et al., 2021). However,
one of the main reasons why NS is calculated per 100 g is standard-
ization. It should be noted that individuals’ portion sizes vary based
on energy/nutritional needs (Fisher et al., 2015) and are also not har-
monized in the EU (Peonides et al., 2022). NS aims to guide healthier
choices through comparisons of similar products in the same food
category, where portion sizes are comparable. Therefore, it provides
meaningful results even if not calculated on a portion (Dreano-Trecant
et al., 2020). A key requirement for an efficient FOPNL is its capacity to
discriminate between products within narrow food categories, and we
have shown that this capacity has been notably improved in NS2023,
especially in categories where limitations related to portion sizes were
highlighted (e.g., vegetable oils, cheeses, and nuts).

In the assessments of the limitations of FOPNLs, it is important to
understand the complexity of distinguishing between healthy and less
healthy foods (Dickie et al., 2022). Although FOPNLs have been found
to be effective in encouraging healthy choices, some of the profiling limi-
tations can never be resolved because the NP of food products only
represents one dimension and cannot consider the complexity of the
diet as a whole (World Health Organization, 2010). However, as has
been shown in many studies, FOPNLs are more effective than only the
back-of-pack nutrition declaration (Rønnow, 2020), and their imple-
mentation is beneficial not just to support healthier food choices but
also as a driver of voluntary food reformulation (Storcksdieck genannt
Bonsmann et al., 2020). The latter effect would be particularly rel-
levant if a harmonized FOPNL system was implemented mandatorily
across the EU marketplace, where the same foods are commonly avail-
able in several countries (Gokani & Garde, 2023). Although FOPNLs
is used particularly by some segments of consumers—more commonly
by those with higher health awareness (Bryla, 2020), the effects of
manufacturers’ food improvements are going much beyond that. Foods
reformulated to get a better FOPNL score are available in the food
supply for all consumers, also for those with low focus on food labels
and healthy diets. This way, it effects of FOPNLs are also relevant for
vulnerable population groups, where highest prevalences of NCDs are
observed (Sommer et al., 2015). Recent study also projected that NS
could on this matter significantly lower annual healthcare spendings
(Devaux et al., 2024).

4.4 Strengths and limitations of the study

Our study presents the first assessment of the improvements to the
NS, a main candidate FOPNL system for adoption across Europe. Main
advantage is that the study builds on a systematic review of all reported
limitations of NS, and that the evaluation was carried out in comparison
to the original (NS2021) and the revised (NS2023) NS profiling algo-
rithms on a large, branded food dataset, covering a wide range of food
products. Furthermore, we used a representative dataset compiled by
a standardized food monitoring study in major grocery stores. However,
limitations of the study should also be noted. Branded food dataset
is nationally representative for Slovenia, with a considerable propor-
tion of imported products (69%). Nevertheless, there are specifics in
the food supply of each country. Considering that our goal was the
comparison of two profiling algorithms, such a limitation could not notably affect the main study outcomes. Another limitation is that for some foods, non-mandatory labeling information was estimated. We estimated %FVL based on ingredient lists that were sometimes limited (missing % each ingredient represents) and fiber content based on calculations from other nutrients and energy. We used established and previously validated methodologies; however, slight deviations from actual values are possible. We should also mention that comparison with dietary guidelines was based on the WHOE profile, which was used as a proxy for FBDG. However, WHOE has been harmonized for Europe and represents a widely accepted profiling model. This model was originally developed for limiting the marketing of unhealthy foods to children and does not incorporate some recent dietary recommendations, such as reducing the intake of red meat and the importance of whole grains (Blomhoff et al., 2023; Willett et al., 2019). Alternative approach would be to use country-specific FBDG (Huybers & Roodenburg, 2023; Konings et al., 2022), which are in Slovenia and not defined to a level that would allow meaningful comparison. Further research should also focus on the assessment of the evolution of NS using country-specific FBDG where available. We should also note that comparison between FBDG and FOPNLs is challenging, as they have very different operating systems and purposes. Therefore, such approach has its own limitations that are also part of our study. We should also note that our literature review focused on all reported limitations of NS algorithm, relevant or not, and that some studies had possible conflicts of interest (see Table S1). We also only focused on the limitations of the NP algorithm and did not cover other dimensions, such as consumer understanding and effect on the food industry and economy.

5 | CONCLUSION

We can conclude that the latest update of NS has addressed the majority of the reported limitations of the NP algorithm, and that we observed an improved, strong agreement with dietary guidelines measured by WHOE. Overall, the updated NS2023 has become a notably stricter profiling model. The resolved limitations mostly addressed Food-based and Component-based (high sugar/salt content, NNS presence) limitations, whereas Other Dimensions of Food Quality remained mostly unaddressed. This study is the first that systematically accessed the revised NS algorithm and was done outside of COEN countries, which were mostly included in the revision of NS, making the study results also relevant to regions where NS is not yet implemented. Discussed limitations highlighted the difficulty of simplifying complex food composition information into FOPNL, which is easily interpretable by consumers. We must mention that the review included all mentioned limitations. Some of the reported limitations may be irrelevant and/or cannot be resolved with nutritional profiling (such as traditional foods and carbon footprints). As frequently highlighted in both media and scientific literature (Besancon et al., 2023), the implementation of the NS is the subject of lobbying from the food industry and other interest groups, some of which finance studies that often generate unfavorable results to NS. Our review also identified studies where potential conflicts of interest might have affected reported limitations of the NS. However, our overall results show that NS2023 is much improved and stricter in comparison to its predecessor, NS2021, and that manipulation of the algorithm is now much more challenging. Nevertheless, only real-life future data can show if NS can affect the food supply and population health on long term. With the implementation of NS2023 on the market, it will be interesting to see how system effects manufacturers and product reformulation. On the other hand, increased strictness of the algorithm could present an even bigger obstacle for its implementation, because of its opponents on the side of the food industry and their supporters. However, the implementation of efficient mandatory FOPNL in the EU is important, and NS has been shown as the most appropriate candidate. Considering that mandatory FOPNL would not only support consumers in healthier food choices but can also stimulate the food industry to actively reformulate processed foods, such policy intervention would also affect consumers who are less interested in healthier diets. The implementation of mandatory FOPNL could therefore also affect the NCD burdens in the most fragile population groups, reducing health inequalities.

AUTHOR CONTRIBUTIONS
Conceptualization; manuscript writing—review and editing: Igor Pravst. Data analyses; formal analyses; writing—original draft preparation: Edvina Hafner. Methodology: Igor Pravst and Edvina Hafner. All the authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGMENTS
Data collections for this study were conducted by the Nutrition Institute (Ljubljana, Slovenia) within the national research program “Nutrition and Public Health” (P3-0395) and the “Infrastructure Programme for the Monitoring of the Composition and Labelling Of Foods” (IO-0054)—funded by the Slovenian Research Agency, and within the Food Nutrition Security Cloud project (FNS-Cloud), which has received funding from the European Union’s Horizon 2020 Research and Innovation programme (H2020-EU.3.2.2.3.—A Sustainable and Competitive Agri-Food Industry) under Grant Agreement No. 863059. The information and views in this report do not necessarily reflect the official opinion or position of the European Union. Neither European Union institutions and bodies nor any person acting on their behalf may be held responsible for any use that may be made of the information contained herein.

CONFLICT OF INTEREST STATEMENT
The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results. I.P. has led and participated in various other research projects in the area of nutrition, public health, and food technology, which were (co)funded by the Slovenian Research Agency, the Ministry of Health of the Republic of Slovenia, the Ministry of Agriculture, Forestry and Food of the Republic of Slovenia, and in the case of specific applied research projects, also by food businesses. Although he has not been
involved in the development or implementation of NS, he was involved in independent studies that assessed NS and disclosed his support for the implementation of mandatory harmonized FOPNL in the EU.

ORCID
Edvina Hafner https://orcid.org/0000-0002-7357-9426
Igor Pravst https://orcid.org/0000-0001-8634-1778

REFERENCES
Besançon, S., Beran, D., & Batal, M. (2023). A study is 21 times more likely to find unfavourable results about the nutrition label Nutri-Score if the authors declare a conflict of interest or the study is funded by the food industry. BMJ Global Health, 8(5), e011720. https://doi.org/10.1136/bmjgh-2023-011720


SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.