Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# The Bias Continuum Framework (BCF) for Medicine–Food Homology: A Derivable Operational Scale across Food, Food-Medicine, and Medicine

**Guanfeng Yang**

jjaarr@163.com

Foshan Yesheng Health Consulting Co., Ltd. https://orcid.org/0009-0009-4933-390X

**Research Article**

**Keywords:** medicine–food homology, catalogue of medicinal and edible substances, nature bias continuum, operational scoring, three-dimensional quantification, four natures of Chinese medicine, evidence tier, blind prediction, falsifiability

**Posted Date:** October 27th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11487275/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

**Additional Declarations:** The authors declare no competing interests.

## Abstract

## Objective

The current Chinese catalogue of medicinal and edible substances applies a "safety–history" criterion that decides whether a substance may be admitted, but it cannot state how strong its cold–hot nature bias is after admission, where its boundaries lie relative to everyday foods and medicines, or how it should be used according to individual constitution. This study constructs an operational scoring framework that uses cold–hot nature bias as a single coordinate axis, is fully rule-disclosed and reproducibly computable, and can be extrapolated to plant-derived substances that lie outside the frozen reference tables.

## Methods

We propose the Bias Continuum Framework (BCF; methodological framework V1), in which foods and medicines are defined as intervals on a continuum of [− 10, + 10], indexed by the nature bias score B (positive = warm/hot, negative = cool/cold, zero = neutral). B is computed from three measurable dimensions—environmental habits (40%), metabolic characteristics (40%), and growth cycle (20%)—under a publicly specified operational scoring rule. Two cut-offs partition the continuum into three zones: $|B|\leq 3$, the low-bias (food-nourishment dominant) zone; $3<|B|\leq 6$, the moderate-bias (medicine–food homology) zone; and $|B|>6$, the high-bias (medicine dominant) zone. The three-dimensional rules and weights were frozen and published on an international preprint platform in 2025, and no rule was retroactively changed for any substance in the present calibration. The 106 national catalogue substances were scored item by item and reported by evidence tier (L1/L2/L3a/L3b/R). Validation followed an ascending staircase: in-system consistency (60 typical foods/herbs), extrapolation of the frozen rules to 19 off-catalogue foodstuffs, and a "predict-before-check" blind prediction on 15 exotic crops and new-resource foods outside the reference tables, supplemented by a cut-off grid sensitivity analysis.

## Results

The 106 substances were classified as 41 low-bias (38.7%), 50 moderate-bias (47.2%), and 15 high-bias (14.2%), a unimodal distribution that bulges in the middle and tapers at both ends. All 30 everyday foods fell in the low-bias zone, and 28 of 30 recognised strongly biased herbs fell in the high-bias zone, with zero food–medicine diagonal misclassification. Of the 19 off-catalogue foodstuffs, 17 were low-bias and 2 (winter melon B = − 3.2, bitter melon B = − 3.3) crossed into the lower edge of the moderate-bias zone. The three-zone structure remained stable for $\alpha\in[2.5, 3.5]$ and $\beta\in[5.5, 6.5]$. B correlated strongly with the rank of Pharmacopoeia natures (Spearman $\rho = 0.910$, $n = 106$, $P < 0.001$; $\rho = 0.938$ in the high-confidence subset), and the plant-derived score span within a single Pharmacopoeia nature label reached about 6.9 points (the four natures govern direction rather than intensity). All 15 high-bias substances were supported by high-confidence evidence (L1/L2); a leave-one-class-out sensitivity analysis that removed all L3b and animal-source (R) entries left the high-bias membership unchanged (0% change). In the blind prediction, 13 of 15 exotic crops agreed in cold–hot direction with public reports; the two exceptions (guarana, cacao) were both rich in caffeine and theobromine (excitatory methylxanthines), precisely exposing a coverage gap of the frozen rule in the metabolic dimension. None of the 15 fell in the high-bias zone, reproducing the general pattern that everyday foods cluster in the low-bias zone.

## Conclusions

The BCF is a reproducible, calibratable, and extrapolatable operational scoring tool that complements rather than replaces the current catalogue: the catalogue governs legality and safety (whether a substance may be admitted or traded), whereas the continuum governs cold–hot displacement and adaptation constraints (how it should be used by constitution, dose, and frequency). The B score measures the direction and intensity of cold–hot bias and is not a direct measure of therapeutic efficacy or toxicity. Because its three-dimensional inputs are botanical measurable parameters independent of classical texts, the framework is, by design, able to compute cold–hot direction and zone membership for exotic crops and new-resource foods "for which botanical data exist but no classical textual record does"; until independent-criterion validation is completed, such results remain "values computed under the frozen rules." The two blind-prediction reversals (the methylxanthine gap) constitute preliminary evidence of the framework's falsifiability. Independent-scorer blind replication (inter-rater reliability) and modern pharmacological criterion association have been registered as subsequent validations.

## 1 Introduction

"Medicine–food homology" is among the most distinctive intellectual legacies of traditional Chinese medicine: food and medicine share a common origin in nature, and a single substance can carry both nutritional and regulatory properties depending on its use, dosage, and combination. Its institutionalised form in contemporary China is the 2002 list of Items Regarded as Both Food and Medicine (87 items) issued by the former Ministry of Health, followed over more than two decades by several expansions prompted by provincial health authorities and supported by safety assessment: six items (Angelica sinensis, Kaempferia galanga, Crocus sativus, Amomum tsao-ko, Curcuma longa, and Piper longum) were added in 2019–2020; nine items (Codonopsis pilosula, Cistanche deserticola, Dendrobium officinale, Panax quinquefolius, Astragalus membranaceus, Ganoderma lucidum, Cornus officinalis, Gastrodia elata, and Eucommia ulmoides leaves) in November 2023; and four items (Rehmannia glutinosa, Ophiopogon japonicus, Asparagus cochinchinensis, and Citrus grandis 'Tomentosa') in August 2024, cumulatively yielding the 106-item Catalogue of Substances Traditionally Regarded as Both Food and Chinese Medicinal Materials (hereafter, the "medicinal and edible catalogue") [1–4]. Substances with dual use have become an important fulcrum for the large-health industry and the sustainable use of Chinese medicinal resources, and their strategic and industrial value has received continued attention from academia and policy-makers [7, 8]. Driven jointly by policy expansion and industry growth, the question of "what counts as medicine–food homology, how it should be judged, and how it should be graded in use" has risen from a purely academic issue to a foundational problem of public health and industrial governance.

The current catalogue system, however, has a structural limitation in its judging logic. Under the 2021 Administrative Provisions on the Catalogue of Substances Traditionally Regarded as Both Food and Chinese Medicinal Materials (Guowei Shipin Fa [2021] No. 36) issued by the National Health Commission and the State Administration for Market Regulation, admission rests mainly on three criteria: a traditional dietary history of at least 30 years, inclusion in the Pharmacopoeia of the People's Republic of China, and toxicological safety centred on acute oral toxicity, three genotoxicity tests, 90-day oral toxicity, and teratogenicity tests [3, 5]. This system is essentially a "safety–history" track: it answers whether a substance is safe for long-term consumption and whether it has a traditional dietary basis, but it does not answer how strong its cold–hot bias is, where its boundaries lie relative to everyday foods and medicines, or at what dose and frequency a given constitution should use it. Two structural phenomena follow. First, catalogue substances are highly heterogeneous along the bias dimension: by the present scale, Chinese yam (B = + 0.21), goji berry (B = + 0.10), and Poria (B = + 0.90) are nearly neutral, indistinguishable from ordinary grains and vegetables, yet appear on the same list as markedly biased substances such as cinnamon (B = + 7.94) and raw Rehmannia (B = − 7.36), which in practice require medicine-level management. More fundamentally, the cold–hot bias of a substance is not a binary "in the catalogue / out of the catalogue" attribute but a continuously distributed intensity spectrum: everyday off-catalogue foodstuffs such as celery, tomato, and cucumber, scored by the same scale, all fall in the low-bias interval $|B|\leq 3$ (respectively − 2.0, + 0.1, and − 2.1), whereas a few items such as bitter melon and winter melon have already crossed slightly below the $|B|=3$ edge (− 3.3 and − 3.2); cases that are "not extreme in bias yet drastic in action" or "not extreme in bias yet mildly toxic" arise among classical medicinal materials (e.g., rhubarb and Asarum). A binary catalogue decision (in = medicine–food homology, out = ordinary food) cannot portray a continuously distributed bias spectrum, and the catalogue system needs an independent, quantifiable bias scale as a complementary dimension.

Quantitative explorations of the four natures of Chinese medicine (cold, hot, warm, cool) have continued for about four decades, yet no unified scale has emerged. Representative approaches include the biothermodynamic work of Xiao Xiaohe and colleagues, who characterised cold–hot medicinal properties using cold/hot plate differentiation and microcalorimetry and proposed a "thermodynamic view of medicinal properties" [11]; the data-driven classification of Wei Guohui et al., who modelled cold–hot properties from similarity of herbal constituents [12]; and the early work of Jiang Xiaobao et al., who proposed cold–hot quantification and a "minimum-quantum principle for cold–hot quantification," addressing the metrological basis of the scale [13]. Reviews likewise note that, over forty-odd years, quantitative studies of medicinal properties remain inconsistent in indicator selection, quantification method, and judging criteria, lacking consensus [18]; on the food side, modern research on the molecular mechanisms of food cold–hot properties (energy metabolism, TRP channels, constituent profiles) is beginning to dialogue with Western functional-food science [19]. Such studies generally rely on experimental assay or constituent retrieval: each substance evaluated requires an independent measurement, there is no common coordinate, and grading standards are not unified. At the same time, reflection on the concept of "medicine–food homology" itself is deepening: Wang Xudong argues that its intellectual origins, conceptual connotation, and extension still require systematic sorting [6], while Shan Feng, Huang Luqi, and colleagues have reviewed its history and development, emphasising the basic division that foods mainly provide nutrition whereas medicines correct bias with bias [7]. Overall, previous quantitative efforts have mostly been "assay-driven," taking case-by-case experimental measurement or constituent retrieval as the unit of evaluation. The present study does not reject this prior work but changes the unit of evaluation: it uses a publicly disclosed, frozen set of rules to compute each food–medicine–homology–medicine item reproducibly and assign it a common coordinate, complementing rather than competing with the existing experimental-assay route.

Existing international and neighbouring work lies on a different dimension from the present framework. In Chinese Herbal Medicines and elsewhere, Yao, Heinrich, Xiao, and colleagues have proposed and advocated a "food–medicine continuum" (FMC), mapping East–West food–medicine relationships in terms of botanical affinity, traditional use and modern constituents, and regulatory evolution, and calling for cross-cultural dialogue [9, 10]. The FMC is a multifactorial taxonomic and cultural framework whose continuum is organised by species category and use tradition; it does not yet provide a single quantifiable, item-by-item reproducible intensity coordinate. As far as the present retrieval reaches, the biothermodynamic approach of Xiao's group [11] and the constituent-similarity modelling of Wei et al. [12] both obtain results item by item through assay or constituent retrieval, and no work has yet placed food, food–medicine, and medicine on a single continuum under one publicly disclosed rule that can be re-applied item by item. The present study positions itself as an open, testable methodological proposal in this direction, not a replacement for existing work.

The relationship between this study and the statutory catalogue deserves explicit statement. From the standpoint of cognitive science, "food" and "medicine" are not classical categories with clear boundaries and necessary-and-sufficient conditions, but prototype categories with a graded structure of "prototypical member–peripheral member–non-member" [14]: rice and Chinese cabbage are prototypes of "food," Aconitum and Coptis are prototypes of "medicine," and a large number of medicinal and edible substances occupy the transitional zone between them. The food–medicine boundary is therefore inherently fuzzy, and the reasonable form of judgement is a continuous interval rather than a one-cut line. On this view, the statutory catalogue can be understood as a set of "administrative sampling points" that the regulator has selected along the whole continuum according to historical and safety procedures—points that span all three zones but are concentrated in the transitional zone (47.2%). What the present study provides is the theoretical coordinate of these sampling points on the nature bias continuum. The two stand in the relationship of "temperature" to "the markings on a thermometer": the catalogue marks a number of certified graduations, while the scale describes the continuous temperature field behind them; neither replaces the other and each corroborates the other. On this basis we propose the Bias Continuum Framework (BCF): food and medicine are defined as intervals on one continuum, indexed by the nature bias score B (− 10 to + 10), with $|B|\leq 3$ the low-bias (food-nourishment dominant) zone, $3<|B|\leq 6$ the moderate-bias (medicine–food homology) zone, and $|B|>6$ the high-bias (medicine dominant) zone; the score is computed from three dimensions (environmental habits, metabolic characteristics, growth cycle) through public parameters and fixed rules, without item-by-item assay. Because all three-dimensional inputs are botanical measurable parameters (habitat moisture, sunshine, photosynthetic pathway, secondary metabolites, growth cycle, etc.), the framework by design does not depend on any classical text—it can not only recompute catalogue substances but also, for exotic crops, new-resource foods, and even newly recorded species "for which botanical data exist but no classical textual record does," compute cold–hot direction and three-zone membership under the frozen rules. This predictive capacity is the key difference from both "assay-driven" and "label back-fill" routes; its positive evidence, applicable boundaries, and a rule blind spot caught in the act by blind prediction are reported in Sections 3.10, 4.7, and 4.8. Using the 106 national catalogue substances as a sample, this study scores them item by item, supplemented by

evidence-tier stratification and staircase validation (with the nature of each validation level honestly labelled), positioning the framework as orthogonal to and complementary with the catalogue system: the catalogue governs "whether a substance may be admitted or traded" (legality and safety), whereas the BCF governs "how it is consumed, and by whom" (cold–hot displacement and adaptation constraints); together they constitute the complete coordinate system for medicine–food homology governance.

## 2 Materials and Methods

# 2.1 The dual-track cognitive model and the positioning of this study

The conceptual premise of this study is the "dual-track cognitive model" proposed in the author's earlier work: human understanding of medicinal and edible substances runs along two tracks. The first is the "safety–history" track, centred on long-term dietary experience and toxicological safety, which answers whether a substance can be eaten and admitted; the second is the "bias–adaptation" track, centred on the cold–hot bias of a substance and its fit with individual constitution, which answers how strong the bias is, for whom it is suitable, and how it is used. The two tracks correspond to different cognitive functions and cannot substitute for one another: a substance can be fully qualified on the safety track (non-toxic under long-term use, with a traditional dietary history) while at the same time showing a marked cold–hot displacement on the bias track that requires restriction for particular constitutions and doses. The full argument of this model is given in the author's prior work [15].

Accordingly, the present study is positioned as an operational scoring tool on the "bias–adaptation" track, not as a replacement for or re-evaluation of the statutory catalogue. A clarification is essential: the B score used here is the operationalised score of the construct "cold–hot nature bias"—that is, the degree of displacement in the cold–hot direction and the corresponding strength of adaptation constraint. It is not therapeutic efficacy, not medicinal potency, and not toxicity. A substance with a high B score is not "more effective"; it only deviates farther from neutrality in the cold–hot direction and is more restricted in applicable population, dose, and frequency. The relationship of B to action drasticity and toxicity is discussed at the end of Section 2.4 and in Section 4.4.

# 2.2 Research materials and data sources

(1) Catalogue substances and list data. The study objects were the current 106 medicinal and edible substances, comprising the first batch (2002, 87 items), the second batch (2019–2020, 6 items), the third batch (2023, 9 items), and the fourth batch (2024, 4 items). The lists and the basis for additions follow the announcements of the former Ministry of Health, the National Health Commission, and the State Administration for Market Regulation, together with the Administrative Provisions on the Catalogue [1–5]. (2) Pharmacopoeia natures and botanical origins. The "nature and flavour" entries under each species in the Pharmacopoeia of the People's Republic of China (2020 Edition, Volume I) served as the traditional directional reference. Compound entries (fresh/dried ginger, raw/processed Rehmannia, bitter/sweet apricot kernel, jujube/sour jujube seed, etc.) were treated under the same principal specification as the B score, and specification differences caused by processing were recorded separately [5]. (3) Three-dimensional ecological parameters and intensity scores. These came from the plant-derived bias scale previously built by the author: three-dimensional intensity scores and threshold rules for 417 common food materials and 570 medicinal materials, together with repeated calibration over 934 samples (485 core, 449 general). The scale and rules were published on an international preprint platform in 2025 [16]; the six animal-derived substances were scored using the animal-derived scale [17]. (4) Item-by-item evidence for the present 106-substance calibration. Where the three-dimensional parameters could be given directly from re-measurable physical quantities (soil moisture, sunshine, photosynthetic pathway, secondary metabolite content, etc.), the entry was labelled L1; where the value already existed in the reference tables it was labelled L2; where species-level recomputation was needed, it was subdivided into L3a and L3b according to whether re-measurable physical parameters were available (see Section 2.7). (5) Blind-prediction samples. To test the framework's predictive power for substances outside the reference tables, a further 15 exotic crops and new-resource foods "with public botanical data but no Chinese classical textual record of nature" were selected; B was computed solely from public botanical/ecological data and checked against public reports of nature or efficacy only after computation was complete (see Section 3.10). All 15 were verified to have no corresponding value in the frozen reference tables (417 foods + 570 medicines).

Rule-freezing statement. All three-dimensional scoring rules, weights (40/40/20), anchors (lean pork B = 0, Aconitum B = + 8.1, Coptis B = − 8.2), and cut-offs (α = 3, β = 6) used in this study were frozen and published with the plant-derived scale on an international preprint platform in 2025 [16]. In the present item-by-item calibration of the 106 catalogue substances and the blind prediction of the 15 exotic crops outside the reference tables, the frozen rules were only applied; no rule, weight, or threshold was retroactively modified for any substance. Hence the 106 results are an application of the frozen rules to a new regulatory object set, and the 15 blind-prediction results are a forward application of the frozen rules to new objects outside the tables, rather than post hoc parameter tuning against the answers. The provenance, freezing, and reproducibility discipline of the rules are detailed in Section 2.5.

# 2.3 The nature bias continuum, anchors, and three-zone cut-offs

The nature bias continuum is defined as a one-dimensional real axis [− 10, + 10], with the nature bias score B as its coordinate: B > 0 denotes a warm/hot bias, B < 0 a cool/cold bias, and B = 0 cold–hot neutrality (neutral nature). Three conventional anchors are set on the axis: lean pork B = 0 (the neutral reference), Aconitum B = + 8.1 (the extreme-hot reference), and Coptis B = − 8.2 (the extreme-cold reference). The anchors are conventional references of the scale rather than physical truths "measured" by experiment; they are analogous to the Celsius scale setting the ice–water mixture at 0°C—an agreed origin and interval that make scores comparable and reproducible. Lean pork was chosen as the neutral anchor because it is consumed in the largest quantity in the everyday diet, is the most neutral in cold–hot nature, and is the most universally used across regions, making it the highest-consensus "neutral" reference. Aconitum and Coptis are respectively the traditional archetypes of "great hot" and "great cold"; values close to but not exceeding the axis ends (± 8.1/±8.2) were chosen to retain scale room for stronger extreme cases.

Two cut-offs divide the continuum into three bias zones: |B|≤3 is the low-bias (food-nourishment dominant) zone, 3<|B|≤6 the moderate-bias (medicine–food homology) zone, and |B|>6 the high-bias (medicine dominant) zone. The cut-offs α = 3 and β = 6 are calibration parameters of the scale, determined under five criteria:  alignment with the natural boundaries of the four-level intensity grading (1–3 weak, 4–5 moderate, 6–7 strong, 8–10 extreme), with 3 the "slight

bias" threshold and 6 the "marked bias" threshold; □ anchor constraints—everyday foodstuffs of the same magnitude as the neutral anchor (e.g., Chinese cabbage, rice) score no more than 3, while recognised strongly biased herbs (Aconitum, Coptis) exceed 6; □ maximising the zone-discrimination consistency of typical foods and typical herbs on the in-system calibration samples; □ ensuring all three zones are non-empty and not grossly disproportionate, so that no zone degenerates; and □ using integer values for ease of regulatory communication and practical use. It must be emphasised that α/β are conventional values fixed on the 2025 preprint calibration samples, not necessary constants derived from a mechanism; the present 106-substance calibration did not adjust the cut-offs, and Section 3.5 tests their neighbourhood stability with a 5×5 α × β grid, reporting their status as "conventional coordinates" rather than "physical thresholds."

## 2.4 Three-dimensional scoring scale: from measurable parameters to the bias score

The B score is computed from a three-dimensional scale. Environmental habits, metabolic characteristics, and growth cycle are chosen because they correspond to the three pathways by which plants develop cold–hot bias in the course of long-term adaptation, and each can be described by botanical measurable parameters without relying on the qualitative judgements of classical Chinese medical texts. The three dimensions are weighted 40%, 40%, and 20%, then summed after multiplication by a direction coefficient (positive for warm, negative for cool):

*B = (environmental-habits intensity score × 40% + metabolic-characteristic intensity score × 40% + growth-cycle intensity score × 20%) × direction coefficient*

The numerical basis for the weights is the stability of repeated calibration in the earlier work: in a pre-experiment of repeated scoring on 100 representative samples, the coefficient of variation (CV) of the 40/40/20 scheme was about 3.2% (SPSS 26.0), smaller than the two alternatives 30/60/10 (about 3.5%) and 50/30/20 (about 4.8%) (Table 5). It must be stressed that this is a selection in the sense of test–retest reliability, not the optimal weight in the sense of predictive validity; the latter would require comparing B with independent criteria (blind scoring, biothermodynamic parameters) across weights, a comparison that has been scheduled for subsequent validation. The 40/40/20 scheme should therefore be understood as a frozen conventional value, not as a proven optimum.

Measurement level of the B score: an ordinal categorical coordinate rather than an interval scale. The current measurement level of B must be stated clearly. Following Stevens' classical distinction of measurement scales [20], what the evidence here most strongly supports is (i) ordinal conclusions (substance A is warmer/cooler than substance B) and (ii) weak interval conclusions (a substance belongs to one of the three ordered categories low/moderate/high). Whether adjacent integer differences of B carry the same cold–hot meaning (interval-scale property) awaits independent-criterion testing; we do not claim that the difference between B = 4 and B = 2 equals that between B = − 2 and B = − 4, nor that B supports strict fractional-value arithmetic. We therefore use B as an ordinal categorical coordinate with conventional graduations: numerical values serve only to display relative position and rank, the three-zone membership and directional ordering are the primary conclusions, and concrete score gaps (e.g., "a span of 6.9 points") should be read as relative spreads on one scale rather than as equidistant physical quantities. This treatment is consistent with the standard requirements of health-measurement instruments regarding measurement properties [21]: interval property, inter-rater reliability, and criterion validity must be established by independent testing and cannot be inferred automatically from the continuity of numerical values. Accordingly, the text reports Spearman rank correlation (ordinal level, primary) together with Pearson correlation (reference only, not an endorsement of interval property), and avoids strong inference from fractional-value arithmetic differences.

The direction coefficients of the three-dimensional weights (xeric, strong-sunshine, positive metabolism indicating warm; hygrophytic, shade-growing, negative metabolism indicating cool), as well as some intensity thresholds (e.g., soil moisture ≤ 20% judged warm, saline-alkali soil pH 8.5–10 judged warm, and the directional assignment of C3/C4/CAM photosynthetic pathways), are not causal laws derived from first principles, but operational scoring rules repeatedly calibrated and then frozen on the earlier reference tables and calibration samples. Their provenance has two layers. The first layer is objective, re-measurable botanical parameters (photosynthetic pathway, specific leaf area, economic maturation period, soil moisture, sunshine, secondary metabolite content), which are independent of Chinese medicine and can be measured by any third party; the second layer is the mapping rule "objective parameters → cold–hot direction and intensity," which is formed by calibration. The strict separation, basis, and reproducibility discipline of the two layers are given in Section 2.5. The three-dimensional scoring elements and their measurement directions are shown in Table 1.

Table 1
Dimensions, weights, and core measurement elements of the three-dimensional scoring scale

| Dimension | Weight | Core measured elements | Warm direction (+) | Cool direction (-) |
|---|---|---|---|---|
| Environmental habits (E) | 40% | Habitat water, soil moisture, sunshine duration and intensity, altitude, soil pH | Arid/semi-arid, low soil moisture, strong sunshine, saline-alkaline soil, high-altitude sun-adapted | Damp-shaded, water-rich, low sunshine, shade-adapted, high-humidity humus |
| Metabolic characteristics (M) | 40% | Photosynthetic pathway (C3/C4/CAM), specific leaf area, content of secondary metabolites (volatile oils / flavonoids / alkaloids) | C4/CAM, high volatile-oil content, positive (warm) secondary metabolism | C3 shade-adapted, large specific leaf area, high cool-related constituents such as flavonoids |
| Growth cycle (G) | 20% | Economic maturity, harvest years, growth rate | Fast-growing, short cycle, annual | Slow-growing, perennial, requiring multi-year harvest (e.g., roots/rhizomes, fleshy roots) |

*Note: The direction coefficient is determined by looking up a table for the specific combination of parameters; special groups (fungi, parasitic plants, high-altitude plants, lichens) have dedicated weight corrections (*Section 2.6*). The complete threshold tables and computation formulas for the three dimensions, and the item-by-item scoring of the 570 medicines and 417 foods, are given in the author's earlier plant-derived scale* [16]. *Every B score corresponds to a verifiable computation chain (three-dimensional inputs → direction coefficient → single-dimension intensity scores → weighting → total score).*

On the distinction between "bias" and "action drasticity/toxicity": B describes only the displacement in the cold–hot direction. A few substances are not extreme in cold–hot bias yet are drastic in action or mildly toxic (e.g., rhubarb and Asarum, B = − 5.80 and + 5.50, both in the moderate-bias zone); their "medicine" attribute is mainly determined by dimensions outside B (dose tolerance window, toxicity, indication constraints). Framework V1 completes only the quantification of the single cold–hot dimension; independent parameterisation of the second dimension, "adaptation-restriction breadth," is left to subsequent work (Section 4.4). It must be particularly stressed that the "secondary metabolites" included in B are limited to constituent classes that are long-term correlated with the cold–hot direction (volatile oils/terpenoids, flavonoids, phenolic acids); strongly toxic alkaloids, cardiac glycosides, and toxalbumins with clear toxicity are not added to the cold–hot B score but belong to the second dimension of "adaptation constraint/toxicity"—a boundary that acquires special significance in the Section 3.10 blind prediction (Sections 4.7, 4.8).

## 2.5 Provenance, freezing, and reproducibility discipline of the rules

To forestall the question of whether the rules were derived from theory or repeatedly tuned against traditional medicinal properties as the standard answer, this section gives a full account of the provenance, layering, and freezing discipline of the rules.

(1) Two-layer provenance. The rules fall into two layers of different nature. The first is the objective-parameter layer: the three-dimensional inputs (photosynthetic pathway, specific leaf area, soil moisture, sunshine, altitude, economic maturation period, secondary metabolite content, etc.) are all botanical measurable physical quantities independent of Chinese medical concepts, measurable by any third party using published methods. The second is the calibration-mapping layer—the mapping rules and weights from "objective parameters to cold–hot direction and intensity"; this layer is not a causal law derived from first principles but an operational rule repeatedly calibrated and then frozen on the earlier reference tables and the 485/934 calibration samples. The accurate description of B is therefore "an operational score based on plant physiological parameters," not "an objective measurement derived from a mechanism"; the framework is an operational instrument, not a causal mechanistic model.

(2) Nature of the mechanistic chain. Several "environment–metabolism–constituent" correlations were observed in the data: soil moisture showed a strong negative correlation with volatile-oil content (about r = − 0.82), and sunshine a strong positive correlation with volatile-oil content (about r = + 0.75); for some substances, the regression coefficients of determination of flavonoid content against the three-dimensional cool score and of gingerol content against the warm score were about $R^2$=0.91 and $R^2$=0.88, respectively. It must be clear that these values are in-sample fit correlations observed over the whole sample set during earlier calibration, not independent leave-one-out cross-validation or measurements on new samples; they are reported here only as empirical observations during rule formation and not as independent validation evidence. The step "constituent → cold–hot efficacy" (volatile-oil constituents act as warm-yang in the body; flavonoids act as heat-clearing) remains a mechanistic hypothesis that awaits modern pharmacological and clinical testing, and we do not treat it as a proven mechanism. In other words, the validity of the BCF rules is tested by predictive validity (consistency with Pharmacopoeia direction, known groups, and external samples; Sections 3.3–3.6 and 3.10), and does not depend on whether the mechanistic chain has been elucidated.

(3) Four-layer structure of the rules. To make it clear to reviewers and users which rules are core and not to be lightly changed and which were added during calibration, all rules are explicitly divided into four layers (Table 2). The exploratory rules (the Pharmacopoeia-nature mapping function and the constituent regression models) are used only for comparison and robustness checks and do not participate in the primary judgement of any substance.

Table 2
The four-layer structure of the BCF rules

| Rule layer | Content | Nature | Role in this study |
|---|---|---|---|
| Core rules | Three dimensions, weights 40/40/20, direction coefficients, the B-score formula | Framework backbone | Not adjusted for any individual substance |
| Calibration rules | Anchors (lean pork 0 / Aconitum + 8.1 / Coptis − 8.2), cut-offs alpha = 3 / beta = 6, per-dimension intensity thresholds | Conventional parameters fixed on calibration samples | Frozen in 2025; not changed here |
| Correction rules | Special-group weights (fungi / parasitic / high-altitude / lichens), conflict arbitration | General corrections for group-level systematic bias | Applied uniformly by group |
| Exploratory rules | Pharmacopoeia-nature mapping function, constituent regression models | Comparison and robustness checks | Do not enter the main judgement |

(4) Discipline of the calibration and freezing periods. The life cycle of the rules has two phases. During the calibration period (before the 2025 preprint), the rules could be revised in light of data, but two iron rules had to be observed: ① revisions could only be made at the rule level (a weight, a threshold, or a group rule), after which all samples had to be recomputed; case-by-case exceptions for a single substance were prohibited; and ② every revision required a general basis (a systematic bias for a class of substances or an interpretable empirical correlation); ad hoc rules designed to make the result of one substance "look good" were forbidden. From the freezing period (publication of the 2025 preprint) onward, the rules, weights, anchors, and cut-offs were fully frozen, and thereafter only applied to any new object (including the present 106 substances, the 15 blind-prediction species, and future substances).

(5) A fixed troubleshooting sequence for anomalous results. When a result computed under the frozen rules conflicts with experience or expectation, direct modification of the rules is not allowed; instead, the following sequence is followed: ① input-parameter check—whether the three-dimensional measured values or assignments are accurate and reliably sourced; ② origin and specification check—whether there is confusion of species, botanical origin, medicinal part, or processing specification (e.g., fresh/dried ginger, raw/processed Rehmannia, bitter/sweet apricot kernel); ③ special-group check—whether a group weight correction should be applied or was applied in error; ④ dimensional explanation—whether the conflict lies in a dimension not covered by B (action drasticity, toxicity, meridian entry; Section 4.4); ⑤ traditional-knowledge re-examination—whether an ancient record needs to be re-examined under contemporary ecological conditions, which must be based on quantifiable contemporary ecological parameter changes (e.g., altitudinal shift of production areas, shortened cultivation cycles, altered soil and sunshine) and supported by data, not solely on the ancient record (Section 4.3); and ⑥ rule version

upgrade—only when counterexamples of the same type are reproducible, accumulable, and point to a systematic defect of the rules is a version upgrade (e.g., v2) initiated, and it must recompute all substances (including already calibrated ones) and publish the change log and the old/new differences. The guarana and cacao cases in Section 3.10 were checked in this sequence: inputs and specifications were correct, the counterexample was reproducible across two species, and both pointed to the same constituent-class blind spot; they therefore constitute a candidate systematic v2 correction (adding a methylxanthine indicator) rather than an ad hoc V1 judgement. The first five steps do not change the frozen rules and operate only at the input, specification, or explanatory level.

(6) Reproducibility and version management. Every B score must correspond one-to-one to a complete computation chain: three-dimensional inputs (with values and sources) → direction coefficient → single-dimension intensity scores → weighting → total. The accompanying electronic supplementary table (evidence-tiered version) retains all inputs, computation processes, and evidence tiers; any user with the frozen rules and input data can mechanically recompute the same score. For L3b literature/expert inference values (Section 2.7), the source materials and scoring basis of the inference must likewise be stated, making them "traceable expert judgements" to be replaced in future by re-measurable parameters. Both rules and data are under version management (this study is V3.3: V3.1 introduced evidence tiers; V3.2 completed item-by-item L3b provenance cleaning, disclosure of overlap with the calibration set, and expansion of external samples, corrected one entry—Cornus officinalis—from + 1.0 to the rule-arbitrated value 0.0, and, under the frozen arbitration rule (environment and metabolism opposite with a difference ≥ 3), re-checked item by item and added three previously omitted arbitrations (mulberry, monk fruit, Chinese olive), so that all 106 scores can be reproduced by the same rule; V3.3 made the predictive and falsifiable character of the framework explicit, added the blind prediction of 15 exotic crops/new-resource foods outside the tables, supplemented the plant sessility–chemical-defense mechanistic hypothesis and the plant–animal bridge, and completed P0/P1 consistency revisions). Every version upgrade is published with the changed content.

## 2.6 Special-group corrections and conflict arbitration

For ordinary terrestrial vascular plants, the core rules are applied directly. For groups with special ecological adaptation mechanisms, where ordinary rules produce systematic bias, unified group weight corrections are set: fungi (e.g., Ganoderma, Poria) do not rely on photosynthetic pathway as their main metabolism, and the environment and metabolism dimensions are reweighted according to their substrate and temperature adaptation; parasitic plants (e.g., Dendrobium officinale) have their habitat soil parameters weakened and host and metabolism dimensions strengthened; high-altitude plants and lichens are given separate weights. These corrections belong to the "correction rules" of Table 2, applied uniformly by group and not to individual substances. When the three dimensions conflict in direction, the frozen conflict-arbitration rule is applied: arbitration is triggered when the environment and metabolism dimensions point in opposite directions and differ in intensity score by ≥ 3, after which the weights are metabolism 60% + environment 30% + growth 10% (metabolism dominant; the basis is that among 50 medicinal plants the agreement rate of the metabolism dimension with traditional natures was about 92%, higher than the about 82% of the environment dimension). In the item-by-item re-check of the present 106, nine entries triggered arbitration: Phyllanthus emblica, Cassia obtusifolia, Chaenomeles speciosa, Nelumbo nucifera seed, Cichorium intybus, Cornus officinalis, Morus alba fruit, Siraitia grosvenorii, and Canarium album. In a validation of 25 direction-conflict cases (8 daily, 17 medicinal), arbitration raised directional agreement from 82% (20/25) to 95% (24/25); this result is from the earlier frozen work [16]. The nature of the 92%/82% and 95% agreement rates must be stated: they all come from within the earlier frozen system, as consistency statistics on the same rules and the same calibration samples (in-sample consistency); the "traditional natures" used in some cases were precisely the tuning basis for the weight arbitration, so they are not an independent criterion and cannot establish external validity for the arbitration rule. The independent validity of the arbitration rule must be reported separately after inter-rater reliability (IRR) and modern pharmacological criterion tests.

The complete arbitration calculation is illustrated with Cornus officinalis. Its three-dimensional inputs are environmental habits E = + 2 (warm), metabolic characteristics M = − 1 (cool), and growth cycle G = 0. Environment and metabolism are opposite, and the intensity difference |+2−(− 1)|=3 reaches the ≥ 3 trigger threshold, so the arbitration weights M60/E30/G10 are used: B = M×0.6 + E×0.3 + G×0.1=(− 1)×0.6+(+ 2)×0.3 + 0×0.1 = − 0.6 + 0.6 + 0.0 = 0.0. The environment and metabolism dimensions exactly offset under the arbitration weights, giving B = 0 and a neutral assignment; its pre-V3.2 value of + 1.0 was corrected to the rule-arbitrated value 0.0 under the frozen rules (version change as described above).

Table 3
Correction directions for special groups

| Group | Why the ordinary rules fail | Correction direction |
|---|---|---|
| Fungi (Ganoderma, Poria, etc.) | No photosynthetic pathway; not distinguished by C3/C4/CAM | Downweight photosynthesis; reconfigure by substrate, temperature adaptation, and metabolism |
| Parasitic plants (Dendrobium officinale, etc.) | Habitat soil parameters do not reflect their adaptation | Downweight soil; strengthen host and metabolic dimensions |
| High-altitude plants | Strong UV and large temperature differences affect secondary metabolism | Strengthen UV and temperature-difference weights |
| Lichens | Fungus-alga symbiosis, not a single plant | Separate weights; calibrated separately |

## 2.7 Evidence-tier stratification

To prevent evidence of different strengths from being treated as equivalent, every B score is tagged with an evidence tier, five in total. L1 is a full three-dimensional computation (all three-dimensional inputs have re-measurable physical quantities); L2 is a value assigned from the reference tables (using the three-dimensional scores of the already frozen scale); L3a is species-level recomputation supported by re-measurable parameters; L3b is species-level recomputation whose inputs rely mainly on integer scoring from literature and expert experience and lack re-measurable physical quantities; R is an animal-derived reference value (using the animal-derived scale, which in this study does not enter the core conclusions with the same weight as the plant-derived scale). The stratification criteria and the count per tier are given in Table 4.

Table 4
Criteria for the evidence tiers

| Tier | Criterion | Re-measurability | Count in this study |
|---|---|---|---|
| L1 | Complete three-dimensional inputs, all with re-measurable physical quantities (soil moisture, sunshine, photosynthesis, constituents, etc.) | Strong | 4 |
| L2 | Adopt three-dimensional scores from the frozen reference tables (417 foods / 570 medicines) | Stronger | 62 |
| L3a | Species-level recomputation supported by re-measurable parameters | Medium | 0 |
| L3b | Species-level recomputation, mainly integer scores from literature records and expert experience | Weak | 34 |
| R | Animal-derived reference value (animal-derived scale) | To be validated | 6 |

*Note: L3a is a reserved tier for upgrading L3b to re-measurable parameters; no entry among the current 106 meets the L3a standard (the 34 species-level recomputations all lack re-measurable physical quantities and are therefore L3b). L3b is the principal source of uncertainty in this calibration; its route sensitivity is reported in* Section 3.7, *and the core conclusion (the high-bias membership) does not depend on L3b.*

## 2.8 Validation design, validation status, and validation protocol

This study sets three layers of validation and honestly labels what each layer can and cannot prove, avoiding any presentation of "internal self-consistency" as "external validity."

(1) In-system consistency test (resubstitution / in-sample consistency). Sixty typical objects (30 recognised everyday foods and 30 recognised strongly biased herbs) were selected, their B scores computed under the frozen rules, and their zone membership checked. It must be honestly stated that the reference-table values of these objects were computed with the same three-dimensional rules, the same weights and thresholds, and that the selection of the typical objects itself drew on traditional knowledge; this is therefore not an independent external-validity test but an in-system consistency test that demonstrates internal self-consistency within the calibration range, without ruling out overfitting or scorer subjectivity. Moreover, the 60 typical objects were selected at the two poles of the continuum (the most typical everyday foods and the most recognised strongly biased herbs), i.e. two-pole (typical-group) sampling; the 96.7% measures two-pole discrimination rather than full-spectrum validity. Pungent-warm everyday foods near the boundary with disputed direction (leek + 0.8, garlic + 1.0, pumpkin + 2.0) still fall in the low-bias zone under the frozen rules; full-spectrum coverage is left to subsequent work.

The degree of overlap between the calibration set and the 106 samples must be further disclosed to delimit the scope of the resubstitution test. Of the 106, the 66 high-confidence substances (L1 4, L2 62) take their values directly from the 417/570 reference tables frozen in 2025 (five involve name correspondences: kunbu = kelp, feizi = torreya seed, huaimi = Sophora flower, jupi = tangerine peel, and zisuzi = perilla seed), which is strict in-sample resubstitution; the 34 plant-derived L3b entries are all outside the frozen reference tables and are new species-level three-dimensional recomputations under the frozen rules (inputs still rely on literature and expert integer scoring, non-blind), of which about five (e.g., Dendrobium officinale) have same-name or related/different-part entries in the 570-medicine table but were in no case directly assigned those values, instead being recomputed in three dimensions by species, with the reference-table entries serving only as comparison; the remaining about 29 have no corresponding entries. The six animal-derived substances are listed separately under the animal-derived scale. Hence the consistency tests and construct-validity correlations here are dominated by in-system resubstitution and cannot be equated with testing on unseen samples; most substances in the new 2023 and 2024 batches (e.g., Cistanche, Ganoderma, Codonopsis, Astragalus, Rehmannia, Ophiopogon, Gastrodia, Panax quinquefolius, Dendrobium) already had values in the 2025 reference tables and cannot serve as natural hold-outs. True external validity must await samples outside the reference tables scored blindly by independent scorers—the 15-crop blind prediction in Section 3.10 is a step in that direction (though not independent-scorer blind, both species and answers are outside the tables).

(2) External extrapolation of the frozen rules. For 19 everyday foodstuffs outside the catalogue (covering vegetables, fungi/algae, fruits, grains, and meats: celery, tomato, bitter melon, cucumber, rice, apple, lean pork, etc.), B was computed on new objects under the frozen rules to demonstrate extrapolation outside the catalogue. It must be honestly stated that the plant-derived foodstuffs (including the newly added rice and apple) still take their values from the 2025 frozen reference tables, while lean pork is the animal-derived neutral anchor; this group therefore expands category coverage (adding grains, fruits, and meats) rather than constituting an independent blind sample outside the frozen tables. It can partly alleviate the concern that the rules "only work for catalogue substances," but it remains exploratory extrapolation; genuine independent extrapolation is borne by the out-of-tables blind prediction in Section 3.10.

(3) Cut-off sensitivity. A 5×5 grid scan was run for $\alpha \in \{2.0, 2.5, 3.0, 3.5, 4.0\}$ and $\beta \in \{5.0, 5.5, 6.0, 6.5, 7.0\}$, reporting the change in three-zone counts and the migration of borderline substances, to test neighbourhood stability. This can show that the cut-offs do not collapse the structure within a small neighbourhood, but it cannot prove that 3/6 is the globally optimal cut-off (the basis for the choice is in Section 2.3).

(4) Direction test of construct validity. Taking the Pharmacopoeia nature direction as the reference for the traditional construct, the Spearman rank correlation between B and the Pharmacopoeia nature rank was computed and reported by stratum (Section 3.6). Because some L3b scores drew on the Pharmacopoeia direction, this correlation carries a certain circularity risk on L3b; the correlation in the high-confidence subset (L1 + L2) and after excluding L3b are therefore reported together.

(5) Validation status and pre-registered subsequent protocol. Validations not yet completed in this study include: □ inter-rater reliability (IRR)—independent scorers who have not seen the existing B values score random samples solely from the frozen rules, computing the intraclass correlation coefficient (ICC) or Cohen's κ to test the repeatability and subjective degree of the rules; □ modern pharmacological criterion association—correlation with independent criteria

such as biothermodynamic parameters (e.g., microcalorimetry) and animal cold–hot models; and □ component measurement of the three-dimensional inputs. These three have been registered as subsequent validation protocols and will be carried out and released in a preprint update once blind scorers and experimental conditions are available. As framework V1, the conclusions should be understood, pending these validations, as "internally self-consistent, directionally consistent, possessing preliminary out-of-tables predictive power, but still awaiting external independent validation."

## 2.9 Scale reliability: earlier repeated calibration

During the earlier construction of the scale, 934 samples (485 core, 449 general) underwent repeated calibration and cross-checking to assess the repeatability of the scale itself. Everyday food materials and corrected medicinal materials reached relatively high agreement rates with traditional natures, the agreement rates of ordinary plants and special groups differed, and among the weight schemes the coefficient of variation of repeated scoring was smallest for the 40/40/20 scheme. This provides earlier evidence for the "repeatability" of the scale, but it remains the repeatability of one scorer within one rule system and is not equivalent to the independent inter-rater reliability described in Section 2.8. The values are given in Table 5.

Table 5
Reliability results of the earlier repeated calibration over 934 samples

| Indicator | Result | Note |
|---|---|---|
| Concordance with traditional natures (everyday foods, core samples) | ~ 98% | 485 core samples |
| Concordance with traditional natures (medicines, after correction) | ~ 96% | After special-group correction and conflict arbitration |
| Concordance of ordinary plants | ~ 97.2% | 350 of 360 food plants; 404 of 421 medicinal plants |
| Concordance of special groups (fungi / parasitic / high-altitude / lichens) | ~ 92.5% | Still lower after group correction |
| Repeated-scoring coefficient of variation, CV (40/40/20) | ~ 3.2% | Pre-experiment on 100 representative samples (SPSS 26.0); better than 30/60/10 (~ 3.5%) and 50/30/20 (~ 4.8%) |

*Note: These are repeated-calibration results (test–retest reliability) from one scorer under the same rule system, not independent inter-rater reliability (IRR). Data are from the earlier frozen work* [16].

## 2.10 Animal-derived and plant-derived: a unified construct with separate operationalisation

The plant-derived three-dimensional scale of this study and the animal-derived three-dimensional scale previously established by the author [17] are two parallel operationalisations of the same bias construct (cold–hot nature bias B) across two major life groups: the construct is unified but the operationalisation is separate, and the two must not be directly mixed in scoring. They are unified in that both operationalise "cold–hot nature bias" as an ordinal score computed by fixed weights from observable ecological/physiological parameters, both adopt the same continuum [− 10,+10], the same neutral anchor (lean pork B = 0), and the same set of three-zone cut-offs; plant-derived and animal-derived scores can therefore be placed side by side, compared, and jointly plotted at the levels of direction and zone membership. They are separate in that animals are mobile and (for livestock) homeothermic, using homeostasis and diet/activity habits as their main adaptation pathways, so their three-dimensional indicators and weights differ from those of immobile plants, which adapt mainly through chemistry and morphology; for example, the animal-derived scale does not use plant-specific indicators such as photosynthetic pathway or volatile-oil content, but instead scores thermoregulation, diet, habitat, and growth cycle. The six animal-derived substances (Table 7 note) are therefore all computed separately under the animal-derived scale [17] and do not enter the core statistics with the same weight as the 100 plant-derived substances; their positions are only marked on the continuum, so that the internal consistency of each scale is not disrupted by cross-group weighting.

From the standpoint of evolutionary proximity, animal-derived substances (especially livestock foods) and humans both belong to the animal kingdom and are closer than plants in material composition, body temperature, and metabolism, and the pathways by which their bias acts on the human body are also more direct; this observation is one methodological reason why animal-derived substances require an independent scale rather than a simple application of the plant rules, but it does not mean animal-derived foods are "milder" or "more biased." Proximity affects the similarity of the action pathway, not the direction or intensity of bias. For the sake of single-paper scope and focus, the animal-derived rules are not expanded here; the complete animal-derived three-dimensional indicators, weights, species calibration, and triple-component formulation synergy system are given in the animal-derived paper [17]. The species-level applicability boundary of the framework is thereby clear: plant-derived substances follow the plant scale and animal-derived substances follow the animal scale, converging on the same continuum, anchors, and cut-offs to form a complete bias coordinate covering both plants and animals.

# 3 Results

## 3.1 Overall distribution across the three zones

Scoring the 106 catalogue substances item by item under the frozen rules gave 41 substances (38.7%) in the low-bias (food-nourishment dominant) zone, 50 (47.2%) in the moderate-bias (medicine–food homology) zone, and 15 (14.2%) in the high-bias (medicine dominant) zone—a unimodal distribution that bulges in the middle and converges at both ends (Fig. 1, Table 6). In direction, 69 were warm-biased, 36 cool-biased, and one (Cornus officinalis) was B = 0; 22 substances were near-neutral with |B|≤0.5. The warmest was cinnamon (B = + 7.94) and the coolest raw Rehmannia (B = − 7.36). The moderate-bias zone

contained the most substances (47.2%), consistent with the institutional intuition that "the catalogue is dominated by the transitional zone"; yet the low-bias zone still accounted for nearly four-tenths and the high-bias zone for 15, showing that the catalogue does not cover only the transitional zone but spans all three zones along the whole continuum (Section 4.1).

The score clustering in the low-bias zone should be noted: 11 substances—smoked plum (wumei), Chinese yam, peach kernel, Platycodon, radish seed, raspberry, adzuki bean, malt, Polygonatum, Codonopsis, and saffron—all had B = + 0.21. This is not the result of independent item-by-item computation but the neutral reference score that the frozen reference tables assign uniformly to a class of foodstuffs "neutral to slightly warm in nature, dominated by starch/polysaccharides, and from moderate habitats"; it reflects the comparatively coarse scoring granularity (minimum increment 0.21) used by that version of the scale in the neutral zone, and this bar is marked in light colour in Fig. 1. It honestly reveals the inadequate resolution of V1 in the low-bias zone: within that zone, B currently mainly supports a binary split of "near-neutral" versus "slightly biased" and is insufficient to distinguish the substances all sitting at + 0.21; after the future introduction of measured secondary metabolites and finer scoring increments, this clustering should be progressively pulled apart. This granularity limitation does not affect the zone judgement or directional ordering in the moderate- and high-bias zones.

Table 6
Distribution of the 106 catalogue substances across the three bias zones

| Bias zone | Threshold | Warm (n) | Cool (n) | Neutral (n) | Total (n) | Proportion |
|---|---|---|---|---|---|---|
| Low bias (food-nourishment dominant) | \|B\| <= 3 | 32 | 8 | 1 | 41 | 38.7% |
| Moderate bias (medicine-food homology) | 3 < \|B\| <= 6 | 28 | 22 | 0 | 50 | 47.2% |
| High bias (medicine dominant) | \|B\| > 6 | 9 | 6 | 0 | 15 | 14.2% |
| Total | - | 69 | 36 | 1 | 106 | 100% |

*Note: Warm-biased (B > 0), cool-biased (B < 0), and neutral (B = 0) are counted by the sign of B; the low-bias zone includes the one neutral substance (Cornus officinalis B = 0, recomputed by rule arbitration*, Section 3.2*); the 22 near-neutral substances (|B|≤0.5) are counted in the low-bias zone. The three-zone total is 41/50/15.*

## 3.2 Evidence-tier results

The distribution of evidence tiers across the three zones is given in Table 7. The key structural fact is that all 15 substances in the high-bias zone are supported by high-confidence evidence (L1 1, L2 14) and include no L3b literature/expert inference values or R animal-source reference values; the low-confidence L3b entries (34) are distributed only across the low-bias (19) and moderate-bias (15) zones, with none entering the high-bias zone. This means that the most regulatorily consequential conclusion of this study—"which substances are most strongly biased and should be managed as medicine-dominant"—rests entirely on re-measurable parameters or the already frozen reference tables, and does not depend on expert inference, whereas L3b mainly affects the fine division between the low and moderate zones. The L1 definition must be clarified: there are four L1 entries in the whole text—Poria (B = + 0.9) and Ganoderma (B = + 0.19) in the low-bias zone, honeysuckle (B = − 5.2) in the moderate-bias zone, and Cistanche deserticola (B = + 6.50) in the high-bias zone (the Table 7 cross-tabulation is 2/1/1). Poria and honeysuckle are conventional full three-dimensional computations; Ganoderma (a fungal group) and Cistanche (a parasitic group) still meet the L1 criterion of "all three-dimensional inputs with re-measurable physical quantities" after the group weight corrections in Section 2.6. The statement in this section that "the high-bias zone has only one L1" is a zone-cross-tabulation count; the other three L1 entries fall in the low and moderate zones, which does not mean the whole-text L1 total is one. It should also be noted that the dimension-by-dimension physical-quantity computation of the L1 anchors was completed in the "computed version of Table 20" of the 2025 frozen system (see the author's series [16]): this study gives a full three-dimensional expansion for Cistanche (parasitic-group correction) in Section 2.6 and the supplementary tables, while Poria, honeysuckle, and Ganoderma directly cite the computed anchor values of the frozen Table 20; their dimension-by-dimension inputs and computation can be checked in [16], and this is a citation of a complete computed result rather than a reference-table value that lacks the computation process.

The complete list of the 15 high-bias substances (each verifiable in Appendix A) is as follows. On the warm side, 9: cinnamon + 7.94, galangal + 7.00, long pepper + 7.00, star anise + 6.80, Agastache + 6.70, black pepper + 6.50, Cistanche deserticola + 6.50 (L1), Amomum tsao-ko + 6.40, and Sichuan pepper + 6.10. On the cool/cold side, 6: raw Rehmannia − 7.36, Ophiopogon − 7.00, dandelion − 6.78, Asparagus (tian dong) − 6.30, mulberry leaf − 6.28, and Houttuynia − 6.28. As a leave-one-class-out sensitivity test, after removing all 34 L3b entries (and the six animal-source R entries) from the sample as a class, the high-bias zone still comprised these 15 substances with a 0% membership change. This means that even if all the least certain evidence is discarded, the most regulatorily consequential conclusion of this study is unchanged in a single character—a stronger moat than the construct correlation coefficient.

Table 7
Cross-tabulation of evidence tier by bias zone

| Evidence tier | Low bias | Moderate bias | High bias | Total |
|---|---|---|---|---|
| L1 full three-dimensional computation | 2 | 1 | 1 | 4 |
| L2 table-assigned | 17 | 31 | 14 | 62 |
| L3a recomputation with parameter support | 0 | 0 | 0 | 0 |
| L3b literature / expert inference | 19 | 15 | 0 | 34 |
| R animal-derived reference | 3 | 3 | 0 | 6 |
| Total | 41 | 50 | 15 | 106 |

*Note: No L3b or R entry enters the high-bias zone; the 15 high-bias substances are L1 + L2 (1 + 14). The six animal-derived substances (honey B = + 0.40, Ejiao/ass-hide glue B = − 2.60, and chicken gizzard lining B = + 1.00 in the low-bias zone; Zaocys B = − 3.50, Agkistrodon B = + 3.60, and oyster B = − 5.50 in the moderate-bias zone) are listed separately under the animal-derived scale and do not enter the core conclusions with the same weight as the 100 plant-derived substances.*

## 3.3 In-system consistency test (with the rhubarb and Asarum counterexamples)

Per Section 2.8(1), this is an in-system consistency test rather than an independent external validation. The confusion matrix for the 60 typical objects is given in Table 8: all 30 recognised everyday foods fell in the low-bias zone (strict agreement 100%); of the 30 recognised strongly biased herbs, 28 fell in the high-bias zone and 2 (rhubarb B = − 5.80, Asarum B = + 5.50) fell in the moderate-bias zone; there was no diagonal misclassification of a food into the high-bias zone or a herb into the low-bias zone. The strict agreement rate was 96.7% (58/60).

Table 8
In-system consistency of the 60 typical objects (confusion matrix)

| Object category | Low bias | Moderate bias | High bias | Subtotal |
|---|---|---|---|---|
| 30 everyday foods | 30 | 0 | 0 | 30 |
| 30 recognised strongly biased herbs | 0 | 2 (rhubarb, Asarum) | 28 | 30 |

*Note: Strict agreement = 58/60 = 96.7%; diagonal misclassification (food → high-bias or herb → low-bias) is 0. The test uses the same rules and is in-system consistency, not independent validity (*Section 2.8*). Two kinds of test must be distinguished: the present test is a two-pole directional diagonal-anchor test, which only checks whether typical foods are misrouted into the high-bias zone and typical strongly biased herbs into the low-bias zone, and does not require B-zone membership to match the eight Pharmacopoeia intensity levels one-to-one. The level differences between Pharmacopoeia label intensity and B-zone membership (e.g., pungent-warm, fragrant and dry substances such as star anise, Agastache, Sichuan pepper, and Amomum tsao-ko enter the high-bias zone under the B scale, while some cool/cold medicinal and edible substances are judged milder than their Pharmacopoeia cold/slightly-cold labels) are intensity-calibration differences, showing a systematic boundary-conservative tendency; they are discussed as the conservatism and directional collapse of the framework in* Table 17 *and* Section 4.6 *and do not constitute directional misclassification.*

The two counterexamples rhubarb and Asarum are not computational errors; rather, they precisely demarcate the measurement boundary of B. Rhubarb is drastic in its purgative action, and Asarum carries the dose warning "not beyond a coin's worth" and mild toxicity; both are unquestionably "medicines" in their attributes. Yet their cold–hot biases are respectively slightly cold and warm, with B = − 5.80 and + 5.50, neither exceeding the |6| high-bias cut-off. This shows that "whether something is a medicine" cannot be decided by the single cold–hot dimension alone: action drasticity, toxicity, the dose tolerance window, and indication constraints form a second dimension independent of cold–hot. In framework V1, B only characterises the cold–hot direction; the medicinal attributes of rhubarb and Asarum are to be explained by the second dimension, "adaptation-restriction breadth" (Section 4.4). The correct reading of the three zones is therefore "strength of cold–hot bias" rather than "whether it is a medicine": substances in the high-bias zone are almost all medicines, but substances in the low-bias zone are not necessarily ordinary foods (Ganoderma and Poria are examples; Section 4.4).

## 3.4 Extrapolation of the frozen rules to off-catalogue foodstuffs

Applying the frozen rules to 19 common foodstuffs not in the catalogue (Table 9, covering vegetables, fungi/algae, fruits, grains, and meats), 17 fell in the low-bias zone, 2 (winter melon B = − 3.2, bitter melon B = − 3.3) crossed into the lower edge of the moderate-bias zone, and none fell in the high-bias zone. Across all 417 common foodstuffs in the earlier reference tables, 414 were low-bias, 3 were at the lower edge of the moderate-bias zone, and 0 were high-bias (99.3%). This shows that the results the frozen rules give on new off-catalogue objects are highly consistent with the prior that "ordinary foodstuffs should be weak in bias," and the framework shows no sample lock-in that would work only for catalogue substances. The crossing of winter melon and bitter melon into the lower moderate-bias edge is consistent with folk experience of them as "cold and cool," unsuitable for those with spleen–stomach deficiency-cold to consume in quantity, indicating that the low–moderate boundary near |B|=3 has practical meaning.

Table 9
Extrapolation results of the frozen rules for 19 off-catalogue foodstuffs (including grains/fruits/meats)

| Zone | Count | Substances (B value) |
|---|---|---|
| Low bias | 17 | Vegetables/fungi-algae: celery − 2.0, tomato + 0.1, cucumber − 2.1, spinach − 1.8, pak choi − 1.2, eggplant − 1.2, luffa − 2.3, pumpkin + 2.0, Chinese chive + 0.8, garlic + 1.0, onion − 0.1, laver − 1.8, kiwifruit − 1.1, shiitake − 1.2; fruit: apple + 0.12; grain: rice (japonica) -0.2; meat: lean pork 0 (animal-derived anchor) |
| Moderate bias | 2 | Winter melon (cannonball type) -3.2; bitter melon (long type) -3.3 |
| High bias | 0 | - |

*Note: Of the 19, 17 are low-bias and 2 (winter melon − 3.2, bitter melon − 3.3) cross into the lower edge of the moderate-bias zone. The original set included kiwifruit (fruit) and shiitake mushroom and laver (fungi/algae), with apple (fruit), rice (grain), and lean pork (meat, the animal-derived neutral anchor) added; animal-derived or fermented/processed products such as egg, milk, vinegar, mutton, and wine are not within the plant-derived scale and are not fabricated here. Plant-derived values still come from the frozen reference tables and constitute category-coverage extrapolation, not an independent blind sample; garlic is scored + 1.0 under the single-bulb (dutou suan) specification. The full 417 common foodstuffs are 414 low-bias / 3 at the lower moderate-bias edge / 0 high-bias (*Section 2.8*).*

## 3.5 Cut-off sensitivity

A 5×5 grid scan of α and β was run (Table 10). With β fixed at 6, as α increased from 2.0 to 4.0, the low-bias zone increased from 35 to 61, the moderate-bias zone decreased from 56 to 30, and the high-bias zone stayed at 15. With α fixed at 3, as β increased from 5.0 to 7.0, the high-bias zone decreased from 37 to 2, the moderate-bias zone increased correspondingly, and the low-bias zone stayed at 41. Within the neighbourhood $\alpha \in [2.5,3.5]$ and $\beta \in [5.5,6.5]$, the three-zone counts changed smoothly without abrupt jumps and the three-zone structure did not collapse; only borderline substances migrated (the 17 substances with $5<|B|<6$ in Section 3.9 sliding between moderate and high, and substances with $|B|\approx 3$ sliding between low and moderate), with no wholesale cross-zone movement. This shows that the 3/6 cut-offs are insensitive within a stable neighbourhood; but as stated in Section 2.3, stability is not global optimality. α/β are conventional coordinates satisfying the five criteria, and the grid results support their positioning as "conventional coordinates" rather than "physical thresholds"—there is no real physical phase transition at the cut-offs, only a communicable boundary needed for management.

Table 10
Grid sensitivity of the cut-offs α and β (three-zone counts)

| Scan condition | Cut-off value | Low bias | Moderate bias | High bias |
|---|---|---|---|---|
| Fix beta = 6, scan alpha | alpha = 2.0 | 35 | 56 | 15 |
| | alpha = 2.5 | 38 | 53 | 15 |
| | alpha = 3.0 (baseline) | 41 | 50 | 15 |
| | alpha = 3.5 | 55 | 36 | 15 |
| | alpha = 4.0 | 61 | 30 | 15 |
| Fix alpha = 3, scan beta | beta = 5.0 | 41 | 28 | 37 |
| | beta = 5.5 | 41 | 38 | 27 |
| | beta = 6.0 (baseline) | 41 | 50 | 15 |
| | beta = 6.5 | 41 | 57 | 8 |
| | beta = 7.0 | 41 | 63 | 2 |

## 3.6 Directional consistency with Pharmacopoeia natures: construct validity and "intensity resolution"

The eight Pharmacopoeia natures were ordered by cold–hot direction (great hot 4, hot 3, warm 2, slightly warm 1, neutral 0, cool − 1, slightly cold − 2, cold − 3); the coding table is Table 11, and compound entries were coded under the same principal specification as B. The correlations of B with this rank are given in Table 12: the full sample gave Spearman $\rho = 0.910$ and Pearson $r = 0.908$ ($n = 106$, $P < 0.001$); after excluding L3b ($n = 72$), $\rho = 0.934$; for the high-confidence L1 + L2 subset only ($n = 66$), $\rho = 0.938$ and Pearson $r = 0.944$. The correlation was in fact higher in the high-confidence subset, indicating that the consistency of B with the traditional cold–hot direction is mainly supported by re-measurable parameters and the reference tables, not by the literature inferences of L3b. This result is an ordinal-level validity test of the "cold–hot direction" construct (whether B arrays along the traditional cold–hot direction), not an independent validation of intensity interval property; per the measurement-level statement in Section 2.4, Spearman rank correlation is the primary basis and Pearson is reference only.

Table 11
Rank coding of Pharmacopoeia natures and the exploratory fixed mapping (comparison only)

| Pharmacopoeia nature | Great hot | Hot | Warm | Slightly warm | Neutral | Cool | Slightly cold | Cold |
|---|---|---|---|---|---|---|---|---|
| Cold-hot rank | + 4 | + 3 | + 2 | + 1 | 0 | -1 | -2 | -3 |
| Fixed mapped B | + 7.82 | + 7.82 | + 5.39 | + 3.00 | + 0.25 | -4.84 | -5.62 | -6.51 |

*Note: Compound entries are coded under the principal specification—ginger takes fresh ginger as slightly warm (+ 1; dried ginger processing to hot is counted separately), apricot kernel takes bitter apricot kernel as slightly warm (+ 1), jujube takes jujube as warm (+ 2), and Rehmannia takes raw Rehmannia as cold (− 3). The fixed mapping function was fitted in-sample from 38 anchors, and its mean absolute error (MAE) on the same-anchor resubstitution was about 0.50; this is an in-training-set resubstitution error without any independent-sample validation and has no independent predictive meaning; it is comparison only and does not enter the primary judgement.*

Table 12
Correlation of B with the Pharmacopoeia nature rank (by stratum)

| Sample | n | Spearman rho | Pearson r (reference only) |
|---|---|---|---|
| Full sample | 106 | 0.910 | 0.908 |
| Excluding L3b (L1 + L2+R) | 72 | 0.934 | 0.937 |
| High-confidence L1 + L2 only | 66 | 0.938 | 0.944 |

*Note: Spearman uses average ranks (ties given average rank), and P values are from a t approximation; all ρ have P < 0.001. Because some L3b scores drew on the Pharmacopoeia direction, the correlation after excluding L3b and in the high-confidence subset are reported together to flag the circularity risk on L3b. Pearson correlation assumes approximately equidistant variables, whereas the interval property of B is not yet established (*Section 2.4*); this column is therefore reference only and is not evidence for interval property or linear intensity.*

If directional correlation only shows that "B is consistent with the direction of the four Pharmacopoeia natures," then within-group dispersion is the framework's irreplaceable contribution—intensity resolution. Under the same Pharmacopoeia nature label, the intensity span of B is very large: among plant-derived substances, the warm group runs from raspberry (about + 0.21) to star anise (+ 6.80), a span of about 6.6 points; the plant-derived neutral group runs from lotus leaf (− 2.2) to liquorice (+ 4.66), a span of about 6.9 points (if animal-derived substances are included, the full-scope neutral span is larger, but the "6.9 points" in this paper always refers to the plant-derived scope). In other words, plant-derived substances all labelled "warm" range from nearly neutral to close to a hot medicine, and plant-derived substances all labelled "neutral" even straddle the low–moderate zones on the B scale. This is direct evidence that the four Pharmacopoeia natures "govern direction but not intensity," and it also shows that back-filling scores directly from labels (e.g., cold → −6.51) would erase real differences of up to 6 points or more among substances sharing a label. The value of the three-dimensional operational score is precisely to re-separate the intensity differences among substances with the same label while keeping direction consistent. It is reiterated that the "6.6/6.9 points" above are relative spreads on one ordinal scale, used to illustrate within-label dispersion, not equidistant physical differences.

## 3.7 Input-route sensitivity of L3b

The three-dimensional inputs of the L3b entries (34) rely mainly on literature and expert experience and are the principal source of uncertainty in this study. To test how sensitive their zone membership is to the "input route," two routes were compared in parallel for these 34 substances: the three-dimensional operational score used here (the main V3 result), and the direct Pharmacopoeia-nature fixed mapping (Table 11). The results are in Table 13: the two routes gave different zone membership for 15 of them (44%); under the three-dimensional route, 0 L3b entries entered the high-bias zone, whereas the Pharmacopoeia mapping route additionally pushed four cold-natured substances (gardenia, mulberry, Lophatherum, and fresh Phragmites root, all mapped to − 6.51) into the high-bias zone.

Table 13
Zone membership of the 34 L3b substances under the two input routes (comparison result of the fourth-layer "exploratory Pharmacopoeia mapping rule" in Table 2)

| Input route | Low bias | Moderate bias | High bias | High-bias members |
|---|---|---|---|---|
| Three-dimensional operational score (V3 main result) | 19 | 15 | 0 | None |
| Pharmacopoeia fixed mapping (comparison) | 12 | 18 | 4 | Gardenia, mulberry, Lophatherum, fresh Phragmites root |

*Note: The two routes differ in zone membership for 15 substances (44%). The mapping route maps the "cold" label uniformly to − 6.51, ignoring intensity differences among substances sharing a label; the three-dimensional route gives weak-bias scores from cultivation, dietary context, and ecological parameters (e.g., mulberry − 3.4, Cassia obtusifolia − 2.3, monk fruit − 2.0, Sterculia/boat-fruited sterculia seed − 2.0).*

Table 13
(continued) Details of the 15 substances whose zone membership flipped (three-dimensional score vs. Pharmacopoeia fixed mapping)

| Substance | Pharmacopoeia nature | Three-dimensional B | Three-dimensional zone | Mapped B | Mapped zone |
|---|---|---|---|---|---|
| Fresh Phragmites root | Cold | -3.80 | Moderate | -6.51 | High |
| Gardenia | Cold | -3.40 | Moderate | -6.51 | High |
| Mulberry | Cold | -3.40 | Moderate | -6.51 | High |
| Lophatherum | Cold | -3.40 | Moderate | -6.51 | High |
| Cassia obtusifolia | Slightly cold | -2.30 | Low | -5.62 | Moderate |
| Monk fruit | Cool | -2.00 | Low | -4.84 | Moderate |
| Sterculia (boat-fruited sterculia seed) | Slightly cold | -2.00 | Low | -5.62 | Moderate |
| Chicory | Slightly cold | -1.30 | Low | -5.62 | Moderate |
| Fermented soybean (dan dou chi) | Cool | -1.00 | Low | -4.84 | Moderate |
| Chaenomeles speciosa (flowering quince) | Warm | + 1.00 | Low | + 5.39 | Moderate |
| Juhong (red tangerine peel) | Warm | + 2.20 | Low | + 5.39 | Moderate |
| Sword bean | Warm | + 2.60 | Low | + 5.39 | Moderate |
| Citrus grandis 'Tomentosa' (pummelo peel) | Warm | + 2.80 | Low | + 5.39 | Moderate |
| White hyacinth bean | Slightly warm | + 3.20 | Moderate | + 3.00 | Low |
| Chinese mosla (xiang ru) | Slightly warm | + 3.40 | Moderate | + 3.00 | Low |

*Note: The three-dimensional B is the main V3.3 result of this study (evidence tier L3b), and the mapped B is the comparison value under the Pharmacopoeia-nature fixed mapping (*Table 11*). Under the three-dimensional route, 0 L3b entries enter the high-bias zone; the mapping route additionally pushes gardenia, mulberry, Lophatherum, and fresh Phragmites root into the high-bias zone, and all four lack re-measurable parameter support.*

Three honest points should be made. First, the Pharmacopoeia mapping route back-fills intensity directly from classical labels, which is itself circular (the label is the conclusion) and erases within-label intensity differences; the three-dimensional score is therefore the main result, with mapped values as comparison only. Second, the three-dimensional intensities of substances such as mulberry, Cassia obtusifolia, Sterculia, monk fruit, gardenia, Lophatherum, and fresh Phragmites root currently rely on L3b expert scoring, and their low–moderate–high assignment is a clear weak point that should be re-checked in future with priority using re-measurable ecological parameters and measured secondary metabolites. Third, and most importantly, the 15 high-bias substances (L1 1 + L2 14) that carry the most regulatory meaning remain unchanged under both routes—the mapping route merely adds four controversial members lacking parameter support and does not overturn any high-confidence core member. The core conclusion therefore does not depend on L3b; L3b route sensitivity only affects the subdivision of the low–moderate zones and a few controversial substances, and the study stratifies them accordingly.

## 3.8 Evolution across catalogue batches

The three-zone distribution by four batches is shown in Fig. 2 and Table 14. Of the 87 items in the first batch, 9 were high-bias (10.3%); of the 19 items in the three subsequent expansions (2019, 2023, 2024), 6 were high-bias (31.6%); of the 4 items added in the fourth batch of 2024 (Rehmannia, Ophiopogon, Asparagus, Citrus grandis 'Tomentosa'), 3 were high-bias (75.0%). It must be emphasised that the above comparison is only a descriptive observation: the later expansion batches are very small (the fourth batch n = 4, and the second and third batches together n = 13), percentages such as 31.6% and 75.0% are strongly affected by individual substances, and the sample size is insufficient for statistical trend inference; this paper therefore does not claim that "the high-bias proportion of later expansions rose significantly," but only records the phenomenon. Qualitatively, this descriptive phenomenon is compatible with the possible evolutionary logic by which the catalogue first admitted substances with uncontroversial safety and the broadest traditional dietary basis, and then, as risk-assessment capacity improved, progressively admitted substances with stronger bias that require more regulated use; but this compatibility is not a statistical proof of that evolutionary logic.

Table 14
Three-zone distribution of the four batches

| Batch | Count | Low bias | Moderate bias | High bias | High-bias proportion |
|---|---|---|---|---|---|
| First batch (2002) | 87 | 35 | 43 | 9 | 10.3% |
| Second batch (2019–2020) | 6 | 1 | 3 | 2 | 33.3% |
| Third batch (2023) | 9 | 4 | 4 | 1 | 11.1% |
| Fourth batch (2024) | 4 | 1 | 0 | 3 | 75.0% |
| Later expansions total | 19 | 6 | 7 | 6 | 31.6% |

*Note: The later expansion batches (especially the fourth batch, n = 4) are extremely small and the high-bias proportion fluctuates strongly; this table is a descriptive record only and does not perform statistical trend tests or significance inference.*

## 3.9 Cut-off boundary substances

Seventeen substances lie at the upper edge of the moderate-bias zone with $5<|B|<6$ (Table 15), the group closest to the high-bias cut-off, including perilla, Angelica sinensis, Astragalus, honeysuckle, fresh ginger, and chrysanthemum. These substances are the most likely to be disputed in positioning between everyday use and medicinal use and best illustrate that "the cut-off is a conventional coordinate rather than a physical threshold": fine-tuning β within [5.5,6.5] makes several of them slide between the moderate and high zones (Section 3.5) but does not change the overall distribution. For these substances, practical management should adopt "zone management" rather than "point-value management": acknowledge that their bias approaches the high-bias level and retain restrictions for particular constitutions and doses, rather than dwelling on which side of the cut-off they happen to lie. Near the low–moderate boundary ($|B|=3$), there is likewise a group of substances with $|B|$ around 2.5–3.5 that are sensitive to fine-tuning of α (e.g., winter melon, bitter melon, mulberry); the same handling principle applies and they are not listed one by one.

Table 15
Cut-off boundary substances (5<|B|<6, in descending order of |score|)

| Substance | B | Evidence | Substance | B | Evidence |
|---|---|---|---|---|---|
| Perilla | + 5.90 | L2 | Mint | -5.52 | L2 |
| Angelica sinensis | + 5.80 | L2 | Ginger (fresh) | + 5.50 | L2 |
| Kudzu root | -5.72 | L2 | Tangerine peel | + 5.50 | L2 |
| Turmeric | + 5.72 | L2 | Oyster | -5.50 | R |
| Long-stamen onion bulb (xie bai) | + 5.70 | L2 | Nutmeg | + 5.50 | L2 |
| Astragalus | + 5.70 | L2 | Eucommia ulmoides leaves | + 5.50 | L2 |
| Honeysuckle | -5.20 | L1 | American ginseng (P. quinquefolius) | -5.16 | L2 |
| Sophora flower bud | -5.10 | L2 | Sophora flower | -5.10 | L2 |
| Chrysanthemum | -5.06 | L2 | - | - | - |

*Note: The 17 substances in this table follow the paper's main criterion 5<|B|<6 (excluding exactly 6.0); 16 are plant-derived and oyster is animal-derived (R, listed separately under the animal-derived scale). Under the electronic-supplementary-table counting criterion 5<|B|≤6 (including exactly 6.0), the count is 22; the extra five are clove, fennel, sharpleaf galangal fruit (yizhiren), and Amomum villosum (all B = + 6.00) and fresh Imperata root (B = − 6.00). The difference between the two criteria arises solely from whether the boundary includes the equals sign, and the membership lists are consistent. Fresh Imperata root B = − 6.00 lies exactly at the upper bound of the moderate-bias zone (|B|=6) and is assigned to the moderate-bias zone; under the main criterion it is not counted here. Dried ginger (B = + 7.42) is a processing specification of fresh ginger and is treated as a processing-slip case, not listed as an independent species.*

**3.10 Blind prediction of exotic crops outside the reference tables: a preliminary test of predictability and falsifiability**

The objects tested in Sections 3.3–3.9, whether inside or outside the catalogue, all take values from or traceable to the 2025 frozen reference tables, which is retrodiction. To test the framework's true predictive power, forward prediction must be made on objects outside the frozen reference tables "whose answers have not been seen." To this end, this section selects 15 representative exotic crops and new-resource foods—all with public botanical/ecological data but no Chinese classical textual record of nature (and not in the frozen reference tables of 417 foods + 570 medicines). B was computed solely from the frozen rules and public botanical parameters, and only after computation was checked against public reports of nature or efficacy.

The blind prediction strictly followed two "predict-before-check" steps to prevent post hoc retuning. Step 1 (prediction): using only public botanical data, the three-dimensional inputs (habitat moisture and sunshine, photosynthetic pathway, secondary metabolites, growth cycle) were determined, and B was computed and assigned to a zone under the frozen rules (ordinary plants E40/M40/G20; when environment and metabolism are opposite and differ in intensity by ≥ 3, arbitration is triggered and M60/E30/G10 used); at this stage no statement about their cold–hot nature was consulted. Step 2 (checking): after the predicted values were fixed, public reports (food nature, traditional medical efficacy, modern functional constituents) were checked to determine whether the direction was consistent. The prediction targets were limited to the two most measurable quantities—cold–hot direction and three-zone membership—without extrapolating meridian entry, specific efficacy, or dose; all results are labelled "values computed under the frozen rules," not verified cold–hot attributes. For "zero-information species" lacking three-dimensional data, no forced score was made; they were left to affinity-based extrapolation and labelled accordingly. All 15 had sufficient data and required no downgrading. The full basis of the three-dimensional scoring, the source URL of each report, and the rationale are given in Appendix B.

Table 18
Blind-prediction results for 15 exotic crops/new-resource foods outside the reference tables (predict-before-check, in descending order of B)

| Species | Family | E | M | G | B | Predicted zone | Public-report comparison (direction) | Judgement |
|---|---|---|---|---|---|---|---|---|
| Maca *(Lepidium meyenii)* | Brassicaceae | 4 | 5 | -1 | + 3.4 | Moderate | Ayurveda warming/heating virya; warming and tonifying, vitality; adaptogen | Concordant |
| Black cumin *(Nigella sativa)* | Ranunculaceae | 3 | 4 | 1 | + 3.0 | Low | Pungent-wind, carminative; black-seed oil warming/dispersing (traditional warm) | Concordant |
| Chia seed *(Salvia hispanica)* | Lamiaceae | 2 | 0 | 1 | + 1.0 | Low | Sweet/neutral, moistening the intestines; no marked cold-hot | Concordant |
| Vanilla *(Vanilla planifolia)* | Orchidaceae | -2 | 3 | -1 | + 0.2 | Low | Slightly warm and harmonising; no strong cold-hot | Concordant |
| Flaxseed *(Linum usitatissimum)* | Linaceae | -1 | 0 | 1 | -0.2 | Low | Pharmacopoeia: sweet, neutral, moistening-dryness (neutral) | Concordant |
| Avocado *(Persea americana)* | Lauraceae | -1 | 0 | -1 | -0.6 | Low | Neutral-moistening, yin-nourishing and fatty; near-neutral | Concordant |
| Noni *(Morinda citrifolia)* | Rubiaceae | 1 | -3 | -1 | -1.0 | Low | Hai ba ji: sour-sweet, neutral, slightly cool | Concordant |
| Cacao *(Theobroma cacao)* | Malvaceae | -1 | -2 | -1 | -1.4 | Low | Traditional/food nature warm, stimulant; rules judged slightly cool (blind spot) | Blind spot / partial falsification |
| Moringa *(Moringa oleifera)* | Moringaceae | 3 | -4 | 1 | -1.4 | Low | Leaves bitter-pungent, heat-clearing and glucose-lowering, folk use as cool (arbitration) | Concordant (arbitration) |
| Artichoke *(Cynara cardunculus var. scolymus)* | Asteraceae | 0 | -4 | -1 | -1.8 | Low | Bitter-cool, choleretic and liver-clearing (Mediterranean cool) | Concordant |
| Guarana *(Paullinia cupana)* | Sapindaceae | -1 | -4 | -1 | -2.2 | Low | Stimulant warm, 4–8% caffeine; rules judged cool (falsification) | Falsified (blind spot) |
| Cyclocarya paliurus | Juglandaceae | -2 | -3 | -1 | -2.2 | Low | Efficacy slightly cool (glucose-lowering, heat-clearing); standards label neutral; partially concordant | Partially concordant |
| Acai *(Euterpe oleracea)* | Arecaceae | -4 | -2 | -1 | -2.6 | Low | Sour-astringent, heat-clearing; cool | Concordant |
| Camu-camu *(Myrciaria dubia)* | Myrtaceae | -4 | -4 | -1 | -3.4 | Moderate | Very high vitamin C, strongly sour, cooling; cool | Concordant |
| Anoectochilus roxburghii | Orchidaceae | -6 | -4 | -1 | -4.2 | Moderate | Clears heat and cools blood; folk records say slightly cold; cool | Concordant |

*Note: E/M/G are the signed scores (− 8 to + 8) of environmental habits, metabolic characteristics, and growth cycle; ordinary weights E40/M40/G20. Moringa triggered conflict arbitration (M60/E30/G10). B is the score computed under the frozen rules, predicting only cold–hot direction and three-zone membership; the "judgement" column follows the public reports checked in Step 2. The full three-dimensional scoring basis, report sources, and falsification rationale are in Appendix B.*

The blind prediction yielded four findings.

First, the directional hit rate was high. Of the 15, 12 were judged directionally concordant (including Moringa, concordant after arbitration) and 1 (Cyclocarya paliurus) partially concordant (the rules judged slightly cool and its heat-clearing efficacy is cool, whereas the sources label it neutral), giving 13/15 consistent in direction with public reports; 2 (cacao, guarana) were inconsistent, a rule blind spot (below). In direction, 4 were warm and 11 cool, with no major warm–cool directional reversal (other than the two excitatory-constituent cases). Second, the zone distribution reproduced "everyday foods clustering in the low-bias zone." All 15 are exotic crops or new-resource foods; 12 fell in the low-bias zone and 3 (maca + 3.4, camu-camu − 3.4, Anoectochilus roxburghii − 4.2) at the lower edge of the moderate-bias zone, with 0 in the high-bias zone. This is fully consistent with the structure of the 106 catalogue substances, the 19 off-catalogue foodstuffs, and even the 417 common foodstuffs—"everyday foods concentrate in the low-bias zone with none high-bias"—showing that the frozen rules do not produce systematic drift toward the high-bias zone on new objects outside the tables.

Third, the conflict-arbitration mechanism is transferable outside the tables. Moringa is warm in habitat (E = + 3, native to tropical dry environments) but cool in metabolic characteristics (M = − 4; its leaves contain glucosinolates and flavonoids and are reported to clear heat and lower blood sugar); the two are opposite with an intensity difference of 7 ≥ 3, triggering the frozen arbitration rule (M60/E30/G10) to give B = − 1.4 (weakly cool), consistent with reports of its leaves "clearing heat, lowering blood sugar, and used folk-wise as cool." This shows that the arbitration rule is not a post hoc patch valid only for in-system calibrated substances but can trigger automatically under unified conditions on new species and yield a direction consistent with reports. Fourth, and most importantly, the blind prediction caught a reproducible rule blind spot in the act. Cacao (B = − 1.4, judged slightly cool) and guarana (B = − 2.2, judged weakly cool) are both warm in public food nature and renowned for stimulation and refreshment (guarana seed caffeine content reaches 4%–8%); the judgements were inconsistent. Both are rich in excitatory methylxanthine alkaloids such as caffeine and theobromine, whereas the "warm constituents" of the frozen rule's M dimension only include volatile oils, terpenoids, and gingerols, omitting methylxanthines—a class of safe but warm/excitatory constituents—causing

systematic under-detection. This is not a toxicity issue (methylxanthines are safe excitatory constituents that should enter the cold–hot B score rather than the second toxicity dimension) but a clear gap in the V1 constituent list; because the counterexample is reproducible on two independent species and jointly points to the same constituent class, following the troubleshooting sequence of Section 2.5(5), it constitutes a candidate systematic V2 correction (adding a methylxanthine indicator in the M dimension) rather than an ad hoc V1 judgement for cacao and guarana.

The nature of this result must be strictly delimited. Section 3.10 demonstrates that "on 15 new objects outside the tables with no classical textual record, the frozen rules can, without changing any parameter, give forward predictions, with most directions consistent with public reports and with a reproducible blind spot actively exposed"; this is positive evidence that the framework is "predictable by design." It cannot yet establish that "predictive validity is established"—the checking was still done by the same scorer, the sample was only 15, and public reports themselves are not a gold standard verified by blind methods. Moreover, the 15 were selected by the researcher under the criterion "with botanical data, without classical records," not randomly sampled, and no sample size or statistical test was preset; the 13/15 directional consistency is therefore only an existence proof of feasibility and does not constitute a statistical inference of the hit rate, whose confidence interval and population extrapolation validity await random expansion. To keep the "predict-before-check" verifiable, the three-dimensional inputs, computed values, and checking conclusions of these 15 were fixed by blockchain timestamping on 21 October 2026 (eBaoquan Electronic Data Preservation Center; record No. 1661064627, verifiable at www.ebaoquan.org by the record number; the right-confirmation file hash and global supervision digest are given in Appendix B and the Data Availability Statement). It should be noted that the timestamp proves that, at the time of deposit, this final version already held in full the stepwise records of "Step 1 predicting solely from botanical data, Step 2 only then checking public reports"; the order of the two steps is additionally guaranteed by the method flow in this section and the split-column structure ("prediction basis / checking conclusion") of the accompanying deposit table, not by the timestamp alone. Independent-scorer blind expansion (IRR), larger samples, and modern pharmacological criterion association are the next rungs of the predictive-validity ladder (Section 4.8).

## 4 Discussion

# 4.1 The statutory catalogue and the nature bias continuum: complementary, not substitutive

The distribution of the 106 substances across the three zones (41/50/15) suggests that the accurate description of "what the catalogue is" should not be "sampling points of the transitional interval"—since 53% of the substances fall in the low- or high-bias zones—but rather: the statutory catalogue is a set of administrative sampling points that the regulator has selected along the whole nature bias continuum according to historical and safety procedures, points concentrated in the transitional zone (47.2%) while also covering both ends. The 41 substances in the low-bias zone (e.g., Chinese yam, Poria, goji berry, honey) correspond to the admission logic of "uncontroversial safety and the broadest traditional dietary basis, therefore admitted first"; the 15 in the high-bias zone come mainly from later expansions, corresponding to the evolutionary logic of "progressively admitting, as risk-assessment capacity improves, more strongly biased substances that require regulated use" (Section 3.8, descriptive observation only).

Hence the catalogue and the BCF stand in the relationship of "temperature" to "the markings on a thermometer": the catalogue marks a number of certified sampling points through safety–history procedures (whether a substance may be admitted or traded), while the BCF describes the continuous cold–hot temperature field behind these points through bias–adaptation rules (how strong the bias is, for whom, and how to use it by constitution, dose, and frequency). The two belong to different tracks and do not substitute for one another. The legality and authority of the catalogue derive from statutory procedure, and the BCF does not endorse the inclusion or removal of any substance; the BCF provides the theoretical coordinate of catalogue substances on the bias dimension and can serve as a technical reference for graded use, public education, and future dynamic adjustment of the catalogue. The BCF does not evaluate the legality or safety of any substance and makes no judgement on catalogue inclusion or exclusion; where the BCF coordinate is inconsistent with catalogue provisions or usage practice, the current laws and regulations always prevail.

A common objection is: dandelion (B = − 6.78), mulberry leaf (B = − 6.28), and Houttuynia (B = − 6.28) can be eaten as wild vegetables or drunk as tea substitutes among the people—why are they judged into the high-bias zone? This is not a framework failure; it precisely shows that between "bias strength" and "how it is actually consumed" lie three mediating variables: dose, frequency, and specification. The same substance used occasionally and in small amounts in cooking (wild vegetable, side dish), regularly over the long term as a tea substitute (daily, continuous drinking), or decocted at a medicinal dose (decoction pieces, extract) represents three different exposure intensities: occasional folk consumption of wild vegetables is mostly low dose, low frequency, and fresh product over a short period, whereas long-term tea substitution or medicinal doses are high-frequency, more concentrated intake, and the actual exposure is not comparable. This corresponds to the track separation in Section 2.1: the catalogue governs "whether it can be eaten," while the BCF indicates "under what dose, frequency, and specification restraint is needed." High bias therefore does not mean prohibited from consumption; it indicates that long-term, high-frequency, large-dose use without pattern differentiation and without a cycle is inadvisable, and in practice one must still combine individual constitution with professional judgement (Section 4.6).

# 4.2 From classical to prototype categories: why a continuous interval is the appropriate form

"Food" and "medicine" are not classical categories with clear necessary-and-sufficient conditions and equal member status, but prototype categories with a graded structure of "prototype members—peripheral members—non-members." Rice and Chinese cabbage are prototypical of "food," while Aconitum and Coptis are prototypical of "medicine," and a large number of medicinal and edible substances occupy the transitional zone between them. The boundaries of a prototype category are inherently fuzzy, and any attempt to make an either-or cut with a single threshold generates disputes at the boundary. The three-zone continuum is the formalisation of this category structure: it does not seek a real physical dividing line between "food" and "medicine," but sets conventional coordinates (α/β) on the continuum for management convenience, and explicitly treats the boundary substances (Section 3.9) as transitional

members "that can slide with fine-tuning of the cut-offs and require zone management." In this sense, "medicine–food homology" is not an either-or administrative identity but a continuous, judgeable property of a substance falling in the transitional interval on the continuum.

## 4.3 The contemporary positioning of classical records: a three-layer partition

The BCF calibration neither evades its relationship with classical records nor treats them simplistically as either "outdated" or "absolute truth." This paper divides the relevant knowledge of substances into three layers (Table 16). The statutory-identity layer (whether in the catalogue, whether it may be traded) is determined by current regulations and is not changed by the BCF. The contemporary-science layer (cold–hot bias, constituents, ecological parameters) is coordinated by the BCF and can be updated as evidence accrues. The conditional-response layer (how well a particular constitution is suited under a particular dose and frequency) depends on the specific conditions of use and requires individualised judgement. The value of classical records lies in providing directional experience and long-term use records; their "direction" component is highly consistent with B ($\rho \approx 0.91–0.94$ at the ordinal level), but their "intensity" component is constrained by historical context and by historical cultivation and use, and must be repositioned at the contemporary-science layer.

Table 16
A three-layer positioning of knowledge about medicinal and edible substances

| Knowledge layer | Content | Basis of judgement | Relationship with the BCF |
|---|---|---|---|
| Statutory-identity layer | Whether in the catalogue; whether it may be traded or produced | Current regulations and catalogue procedure | Not changed by the BCF; only a complementary coordinate |
| Contemporary-science layer | Cold-hot bias, constituents, ecological parameters | Three-dimensional operational score / measurement | The BCF provides the B score and updates as evidence accrues |
| Conditional-response layer | Suitability under a particular constitution, dose, and frequency | Individualised use and clinical judgement | The BCF indicates the strength of restriction; it does not replace judgement |

## 4.4 Future extension: from one-dimensional bias to a two-dimensional adaptation space

B only quantifies the cold–hot direction, and the counterexamples of Ganoderma and Poria clearly expose the inadequacy of a single dimension. Ganoderma B = + 0.19 and Poria B = + 0.90 are both in the low-bias zone, yet in traditional cognition they are medicines rather than everyday foods; both labelled "neutral," pork is an ordinary food while Ganoderma is a medicine. The difference lies not in cold–hot bias but in another set of dimensions—action drasticity, toxicity grade, the latitude of daily dose, and the breadth of constitution constraints. This paper collectively calls this the "adaptation-restriction breadth": cold–hot bias answers "how far a substance deviates in the cold–hot direction," while adaptation restriction answers "by how many people, in what dose, and at what frequency it can be used in daily life." Rhubarb and Asarum (not extreme in cold–hot bias but drastic in action or mildly toxic) are the converse examples of the same gap. It must be emphasised that the "adaptation-restriction/toxicity" dimension takes in strongly toxic alkaloids, cardiac glycosides, and toxic proteins with definite toxicity; the methylxanthines such as caffeine and theobromine found in Section 3.10, by contrast, are safe excitatory constituents that should have entered the cold–hot B score but were missed by V1. The two are different in nature and must not be conflated (Section 4.7).

A complete food–medicine determination should therefore be a two-dimensional space, with the B score (cold–hot bias) as the horizontal axis and adaptation-restriction breadth as the vertical axis. However, this study did not independently compute adaptation-restriction breadth for the 106 substances, so the paper only proposes it as a conceptual and methodological outlook and does not fabricate two-dimensional empirical results. Independent parameterisation of the second dimension (measurable proxies for toxicity grade, dose window, and constitution constraints) is a clear direction for future work; until then, substances such as Ganoderma and Poria that are "low in cold–hot bias but high in medicinal restriction" should be understood as "low bias does not necessarily mean an ordinary food," and their statutory identity remains that set by the catalogue. It must also be made clear that V1 only characterises the "substance" side of the cold–hot attribute and does not at all cover the "individual constitution adaptation" side: the same warm substance may tonify a yang-deficient person but add fuel to fire for a yin-deficient person. Constitution-based pattern matching is the core task of V2, and readers should not map themselves onto a single B score.

## 4.5 Limitations

The limitations of this study should be stated honestly. First, B is an operational score based on plant physiological parameters, not a causal mechanism model: the "objective-parameter layer" is re-measurable, but the calibrated "parameter→cold–hot direction" mapping layer was selected and frozen on 485/934 samples, and its mechanism chain (volatile oils→warm yang, flavonoids→clearing heat) awaits modern pharmacological and clinical testing (Section 4.7 offers a testable hypothesis, not a proven cause). Second, the validation strength is limited: the test on the 60 typical objects is in-system consistency (resubstitution), not independent external validity; although the 15 blind predictions in Section 3.10 are outside the reference tables with neither species nor answers participating in calibration, the checking was still done by the same scorer and is not yet independent-scorer blind work. Inter-rater reliability (IRR), modern-pharmacology criterion association, and constituent measurements are not yet completed and are listed as pre-registered follow-up protocols. Third, the three-dimensional inputs of the 34 L3b substances rely on literature and expert integer scoring, and their low–moderate zone assignment is route-sensitive (Section 3.7), a priority for future measurement-based rechecking. Fourth, framework V1 quantifies only the single dimension of cold–hot bias; the adaptation-restriction breadth dimension is not yet parameterised.

Fifth, the six animal-derived substances use an independent scale and a small sample, and are listed separately here rather than given the same weight as plant-derived substances. Sixth, this is a methodological framework with no animal, cell, or constituent measurement; its positioning is to propose a reproducible, calibratable instrument and to complete a small-sample calibration plus forward blind prediction, not to provide high-grade evidence-based evidence. Seventh, B currently supports only ordinal and weak-zone conclusions, its interval property is unverified (Section 2.4), and the low-bias zone has the granularity problem of 11 substances clustering at + 0.21, indicating insufficient resolution within the low-bias zone in V1; point-value arithmetic and within-

group "score differences" are therefore to be read only as relative spreads. Eighth, the framework shows an identifiable conservative tendency at the boundaries: whenever a substance has disputed direction or conflicting evidence (e.g., Cornus officinalis after arbitration to B = 0, and several arbitration-triggered entries), the result mostly collapses toward the neutral zone, showing a "rather judge neutral than judge extreme" pattern. This conservatism lowers the false-alarm risk for strongly biased substances (consistent with the 0% L3b in the high-bias zone), but it may also systematically underestimate the true bias of some boundary substances; moreover, the methylxanthine blind spot exposed by the blind prediction (cacao, guarana) shows that V1 has a same-direction omission for "warm/excitatory constituents other than volatile oils and terpenoids." Whether these two boundary tendencies (directional collapse and constituent blind spot) are systematic and how large they are must be tested in V2 with IRR and constituent measurements; this paper neither evades nor patches them in advance.

## 4.6 Use-intensity outlook (to be enabled after V2 parameterisation)

To help readers understand the possible practical meaning of B, Table 17 gives a use-intensity outlook by zone. It must be stated that, at the V1 stage, this table is only an illustration to help understand the meaning of the zones; it is not a dosage regimen, not a judgement of efficacy or toxicity, and does not constitute any form of regulatory translation or use recommendation. Until the second dimension "adaptation-restriction breadth" is independently parameterised and B passes independent validity validation, this table is not recommended for actual edible, medicinal, or product-labelling decisions. Any specific use must follow a physician's advice. Only after V2 completes the parameterisation of the second dimension can this table be upgraded from an "illustrative outlook" to an "operational use-intensity reference."

Table 17
Use-intensity outlook by B-score zone (V1 illustration, not yet enabled; not medical advice)

| **\|B\| range** | **Example positioning** | **Use intensity (illustration only; not a recommendation)** |
|---|---|---|
| \|B\| <= 2 (near-neutral to weak) | Everyday tea substitutes, home dishes | May be used routinely and in rotation |
| 2 < \|B\| <= 3 (upper edge of low bias) | Food-nourishment ingredients | Intermittent, rotated use advised; not long-term single-item large amounts |
| 3 < \|B\| <= 6 (moderate bias) | Medicine-food homology conditioning | Use by cycle and by constitution advised; consult a professional when needed |
| \|B\| > 6 (high bias) | Medicine dominant | Pattern-differentiated use advised; self-directed daily or long-term use not recommended |

*Note: At the V1 stage this table is only an illustration of zone meaning, to be enabled after the second dimension is parameterised and independently validated. The B score is not a medication dose, not toxicity, and not efficacy; this table does not constitute edible, medicinal, or therapeutic advice. V1 only characterises the cold–hot bias of the "substance" and does not include the "individual constitution adaptation" dimension—the same warm substance has opposite meanings for yang-deficient and yin-deficient persons. Special populations, pregnant women, children, and patients should consult a licensed physician.*

## 4.7 Plant sessility–chemical defense: a testable mechanism hypothesis

Why can the three-dimensional rules predict cold–hot direction? This section offers a testable mechanism hypothesis that explains, rather than replaces, the frozen rules. Its core logic is: plants cannot move and cannot evade stress by migration, nest-building, or thermoregulation, so over long evolution they invested their survival strategies mainly in chemical defense and morphological adaptation. This is the fundamental reason plant secondary metabolites are far more abundant than animal ones, and it is the material basis of the plant "cold–hot bias." Habitat pressures (water, light, temperature, salinity, pathogen and herbivore pressure) shape the plant secondary-metabolite profile through two interrelated pathways: the first is direct physiological adaptation, and the second is the growth–defense trade-off—the harsher the environment and the fewer resources available for rapid growth, the more plants tend to invest carbon and nitrogen resources in defensive secondary metabolism.

On this basis the framework makes falsifiable directional predictions for three habitat types.  Arid, strong-sunshine, large-temperature-range, high-altitude, or saline–alkaline habitats (e.g., the tropical xerophytes cinnamon, dried ginger, long pepper, and Sichuan pepper) force plants to enrich lipophilic, warm, stimulating, cold-dispelling excitatory constituents such as volatile oils, terpenoids, gingerols, and amides, corresponding three-dimensionally to high positive E and M scores and predicting a warm bias; these constituents mostly act in the human body to dilate vessels, promote heat production, and excite metabolism.  Damp-shaded, low-light, water-rich, humus-rich habitats (e.g., Rehmannia, Ophiopogon, dandelion, Houttuynia, and Anoectochilus roxburghii) lead plants to enrich cool, astringent, heat-clearing constituents such as flavonoids, polyphenols, organic acids, bitter glycosides, and polysaccharides, corresponding to negative E and M scores and predicting a cool bias; these constituents mostly act in the body as antioxidant, anti-inflammatory, and overactive-metabolism-suppressing.  The growth cycle (G dimension) reflects the rhythm of defense investment: fast-growing, short-cycle plants have no time to accumulate large quantities of secondary metabolites and tend to be mild, whereas perennial, slow-growing root and rhizome species requiring several years before harvest have longer to accumulate defensive constituents and tend to be stronger (consistent with the experience that many perennial root/rhizome herbs are strongly biased). This hypothesis is consistent with the growth–defense trade-off of modern phytochemistry and with the direction of molecular-mechanism studies of food cold–hot natures [19], linking "habitat–metabolism–constituent–human cold–hot" into a causal chain testable at every link.

Two classes of secondary metabolites must be strictly distinguished; this is the boundary of the hypothesis. The first class is cold–hot-related constituents—volatile oils/terpenoids, gingerols/amides, flavonoids, phenolic acids, organic acids, polysaccharides—which are safe at ordinary edible doses and whose cold–hot direction can enter the B score. The second class is toxicity-related constituents—strongly toxic alkaloids, cardiac glycosides, toxic proteins, and lectins with definite toxicity—which characterise the "safety window/adaptation restriction," belong to the second dimension (Section 4.4), and are not

counted into the cold–hot B score simply for being "strong." A constituent being "strong" is not equivalent to "warm" or "cool," nor to "toxic"; it must be assigned to the appropriate dimension by its mode of action in the human body. The methylxanthines of Section 3.10 (caffeine, theobromine) fall at a meaningful position: they are safe, non-toxic excitatory constituents whose food nature clearly points to warm excitation (refreshment, promoting metabolism, heat production); they should have entered the M dimension of the cold–hot B score like the volatile oils, but were omitted from the V1 frozen constituent list. This is not a problem of mistaking a toxic constituent for a cold–hot one, but an incompleteness of the cold–hot constituent list itself; it is therefore a concrete gap that can be tested immediately and falsified or confirmed by a single constituent-quantification study.

Finally, the status of this hypothesis must be made clear. It is a falsifiable hypothesis, not a proven causal conclusion; the rule validity of the BCF does not depend on the hypothesis being verified first—on the contrary, the hypothesis provides a testable target for future constituent measurement and pharmacological experiments. Directly testable predictions include: after controlling for species affinity, habitat aridity/sunshine intensity should correlate positively with volatile-oil and gingerol content and with ex vivo/in vivo heat-production and vasodilation indicators; damp-shaded habitats should correlate positively with flavonoid, polyphenol, and organic-acid content and with anti-inflammatory, antioxidant, and overactive-metabolism-suppressing indicators; after adding a methylxanthine indicator, cacao and guarana should return from "judged cool" to "slightly warm/warm," with the return magnitude proportional to their caffeine/theobromine content. If any prediction is refuted by systematic evidence, the corresponding version revision should be triggered per the rule in Section 2.5(5).

## 4.8 The framework's prediction scope, boundaries, and the three-level validation ladder

Synthesising the blind prediction of Section 3.10, this section specifies the applicable prediction scope of the BCF plant-derived scale. Applicable objects include:  exotic crops with botanical data but no classical records (e.g., maca, chia seed, acai, camu-camu);  new-food raw materials / new-resource foods (e.g., Moringa leaves, Cyclocarya paliurus, Anoectochilus roxburghii);  new cultivated varieties, newly introduced or newly recorded plant species;  existing species cultivated under new ecological conditions with quantifiably changed ecological parameters. The common feature of these objects is that the three-dimensional inputs (habitat, photosynthesis, secondary metabolism, cycle) can be described from public botanical data, so the cold–hot direction and three-zone membership can be computed under the frozen rules without waiting for Pharmacopoeia inclusion or catalogue admission. This is the core difference of this framework from the "measurement-driven" and "label-back-filling" routes: it is a ruler that can be read before the answer is seen.

At the same time, the cases that do not apply or require downgrading must be specified:  for "zero-information species" with completely missing three-dimensional parameters, one must not guess from name or morphology; they should be downgraded to affinity-based extrapolation, clearly labelled, and scored only after data are complete;  animal-derived substances (livestock, aquatic products, milk, eggs, etc.) use the animal-derived scale [17] and do not apply plant photosynthesis/volatile-oil indicators;  mineral medicines, fermented and processed products (e.g., vinegar, wine, soy sauce, fermented tea) have their bias determined by processing rather than living ecology and are outside the plant-derived three-dimensional scale;  even where a cold–hot score can be computed for strongly toxic substances (containing strongly toxic alkaloids, cardiac glycosides, etc.), it cannot represent their use safety and must be handled by the second adaptation-restriction dimension;  affinity extrapolation may only be used for congeneric species with similar ecological parameters and must be labelled as extrapolated. For objects outside these boundaries, the framework should not give seemingly precise scores.

Validation of the framework proceeds along a three-level ladder of increasing independence (Fig. 3). The first level is retrodiction (the 106 catalogue substances): the objects are mostly within the frozen reference tables or the calibration scope of the rules, demonstrating internal consistency of the rules on known objects, with the lowest independence. The second level is extrapolation (the 19 off-catalogue everyday foodstuffs): the objects are outside the catalogue but their values still come from the frozen reference tables, expanding category coverage, with medium independence. The third level is blind prediction (the 15 off-table exotic crops/new-resource foods): both species and answers are outside the reference tables, with prediction preceding checking, the first forward-predictive test on objects "not seen before," with the highest independence, though still done by the same scorer. Up the ladder, the results are ever less suspect of "seeing the answer before tuning the rules," and the evidence becomes ever more independent.

Two further rungs must be climbed above this ladder before "predictive validity" can be established: the fourth level is independent-scorer blind expansion (IRR), in which scorers who have never seen the existing B values score random samples independently under the frozen rules alone, with ICC/κ computed; the fifth level is modern-pharmacology criterion association (biothermodynamics, animal cold–hot models, constituent measurement). This paper completes through the third level, so its conclusions must strictly distinguish two propositions: what may be claimed is that "the framework is predictable by design—the rules do not depend on classical records, forward computation can be made for new off-table objects, and of the 15 blind predictions 13 were directionally concordant and a reproducible blind spot was actively exposed"; what may not yet be claimed is that "predictive validity is established—that requires the independent blind work and criterion evidence of the fourth and fifth levels." This distinction is both a requirement of academic honesty and an expression of the framework's falsifiability: a prediction system that explicitly states where it will err and fixes its errors by blockchain timestamping is one qualified to be tested and iterated by the scientific community.

## 5 Conclusions

This paper proposes the Bias Continuum Framework (BCF), which generates the bias score B from three measurable dimensions—environmental habits, metabolic characteristics, and growth cycle—under frozen operational scoring rules, and unifies food, medicine–food homology, and medicine into the low, moderate, and high bias zones of a [− 10,+10] continuum. Item-by-item calibration of the 106 national catalogue substances yielded a 41/50/15 distribution (38.7%/47.2%/14.2%), bulging in the middle and converging at both ends. The in-system consistency test showed no food–medicine diagonal misclassification; the frozen rules placed over 99% of off-catalogue foodstuffs in the low-bias zone; the cut-offs did not collapse the structure within the stable neighbourhood; B was highly consistent with the Pharmacopoeia cold–hot direction (Spearman ρ = 0.910–0.938 at the ordinal level); the 15 most regulatorily consequential high-bias substances were all supported by high-confidence evidence; and after removing all L3b entries the high-bias membership change rate was 0%.

More importantly, because all three-dimensional inputs are measurable botanical parameters independent of classical records, the framework possesses forward-predictive capability by design. The blind prediction of 15 exotic crops/new-resource foods "with botanical data but no classical records" (predict-before-check, blockchain-timestamped, record No. 1661064627) showed that 13/15 were directionally consistent with public reports and the zone distribution reproduced the "everyday foods clustering in the low-bias zone"; the Moringa case demonstrated that the conflict-arbitration rule is transferable outside the tables; the cacao and guarana cases caught in the act a reproducible blind spot of the frozen rules toward methylxanthine warm/excitatory constituents, a clear candidate for a V2 added indicator. This paper also offers the testable "plant sessility–chemical defense" mechanism hypothesis and bridges the animal-derived scale by "unified construct, separately operationalised." The BCF is therefore not a measurement report on 106 substances, but a general ruler marked with the uncertainty of its graduations yet able to be read before the answer is seen: its rules were frozen in 2025 and are reproducible item by item; this paper completes through the third level of the validation ladder (blind prediction), with inter-rater reliability and modern-pharmacology criterion association pre-registered as follow-up validation. The BCF is complementary to the current catalogue—the catalogue governs legality and safety, the BCF governs cold–hot deviation and adaptation restriction; the B score measures cold–hot bias rather than efficacy or toxicity. It is hoped that this coordinate provides a testable, falsifiable, and iterable quantitative starting point for graded use of medicine–food homology, public education, dynamic catalogue management, and the prospective cold–hot assessment of exotic crops and new-resource foods.

## Declarations

**Conflicts of interest**

The author, Guanfeng Yang, is the legal representative of Foshan Yesheng Health Consulting Co., Ltd. and the seventh-generation inheritor of "Yeshengtang" (founded in 1816); the company is engaged in the development, original-equipment manufacturing (OEM), and operation of food-nourishment soup products, health-preserving dew wines, and related products. The author discloses this commercial affiliation in good faith: the design, data calibration, analysis, and formation of conclusions of this study were completed independently, without funding or interference from any commercial organisation; the results do not constitute an efficacy or health claim for any specific product or commodity, nor are they used to endorse the market sales of any particular product. The bias score (B) proposed here is an academic operational score and does not represent efficacy, medicinal potency, or toxicity.

# Ethics statement

This is a theoretical and methodological study based on public botanical data, regulatory documents, and published literature; it involves no human subjects, human materials, or personal data, and no vertebrate experiments or animal-tissue collection, so ethics review and informed consent were not required. The medicinal and edible substances involved are all legally circulating statutory substances, and the study includes no edible, medicinal, or interventional trials.

**Compliance and use statement**

This is an academic theoretical framework and methodological discussion; it does not replace any statutory catalogue, standard, or diagnostic–treatment guideline, and it does not constitute advice on the edible, medicinal, or therapeutic use of any specific substance. The statutory identity, permitted scope of use, and production-and-operation requirements of the substances mentioned are governed in all cases by current laws and regulations, the Catalogue of Substances Traditionally Regarded as Both Food and Chinese Medicinal Materials, and the relevant announcements. Special populations, particular constitutions, and those with diseases should consult a qualified licensed physician or professional before using any medicinal and edible substance. The B score is an operational score of cold–hot bias; it does not represent efficacy, medicinal potency, or toxicity, nor is it used to judge the legality, safety, or catalogue membership of a substance. If the framework's computed results are inconsistent with current laws, regulations, standards, or catalogue provisions, the laws, regulations, and standards always prevail. This paper does not constitute any basis for diagnosis, treatment, medication, or health-food efficacy claims.

**Statement on the blind-prediction exotic crops and new-resource foods in Section 3.10**

The B values listed are only "values computed under the frozen rules," used to test the framework's predictive power; they do not mean that the relevant substances have obtained legal status in China as food or medicinal and edible substances, nor do they constitute a judgement on their edible safety, efficacy, or market access. The use of any new-food raw material is subject in all cases to the safety review and announcements of the national health authorities.

# Funding

This study received no specific grant funding from any public, commercial, or not-for-profit sector.

# Author contributions

Guanfeng Yang: research conception and framework design; construction of the three-dimensional scoring rules and data system; calibration and analysis of the 106 substances and the 15 blind-prediction substances; drafting and finalising the manuscript; responsibility for the entire text. The author thanks the authors of the cited literature and the providers of public materials for facilitating the literature search.

# Data availability statement

The bias scores, three-dimensional computation process, evidence-tier labels, and validation data supporting the conclusions of this paper are provided with the paper as an electronic supplementary table (Excel, V3.3); the three-dimensional inputs, computed values, public-report comparison, and judgement rationale for the 15 blind-prediction species in Section 3.10, together with the corresponding source URLs, are given in Appendix B and in the added "off-table blind prediction" worksheet of the accompanying electronic supplementary table. These data are planned to be uploaded to an open-data repository (e.g., Zenodo) to obtain an independent dataset DOI.The three-dimensional inputs, computed values, and checking conclusions of the 15 blind predictions were fixed by blockchain timestamping on 21 October 2026 (eBaoquan Electronic Data Preservation Center, record No. 1661064627; right-confirmation file hash 632940aa3142869732dfab160887a98a039ccc5df48e71d5a6415cc609d2c25477b9d0c7a2e2eb67a2b5c49879cda76df9d80ccc9f37c8d8db9d50c07302bd1 global supervision digest f7b2b9e14f116667bd3897afcf5fc3f9951eafee; verifiable at www.ebaoquan.org), fixing the timestamp held by this final version; the order of the two "predict-before-check" steps is guaranteed by the method flow in Section 3.10 and the split-column structure of the deposit table. The three-dimensional scoring rules and the 485/934 repeated calibrations are detailed in the author's series [16–17]. This paper is framework V1; inter-rater reliability (IRR), modern-pharmacology criterion association, and constituent measurement are pre-registered follow-up validations, for which an updated preprint version will be released after completion.

## References


1. Ministry of Health of the People's Republic of China (2002) Notice on further standardising the administration of raw materials for health foods (Wei Fa Jian Fa [2002] No. 51) [Z]. Ministry of Health, Beijing. [in Chinese]
2. National Health Commission of the People's Republic of China (2019) Announcement on six newly added substances traditionally regarded as both food and Chinese medicinal materials, including Angelica sinensis (Announcement No. 8 of 2019) [Z]. National Health Commission, Beijing. [in Chinese]
3. National Health Commission of the People's Republic of China; State Administration for Market Regulation. Provisions on the administration of the catalogue of substances traditionally regarded as both food and Chinese medicinal materials (Guo Wei Shi Pin Fa [2021] No. 36) [Z] (2021) [in Chinese]
4. National Health Commission of the People's Republic of China Announcement on nine newly added substances traditionally regarded as both food and Chinese medicinal materials, including Codonopsis (Announcement No. 9 of 2023); Announcement on four newly added substances traditionally regarded as both food and Chinese medicinal materials, including Rehmannia (Announcement No. 4 of 2024) [Z]. 2023–2024. [in Chinese]
5. Chinese Pharmacopoeia Commission. Pharmacopoeia of the People's Republic of China: Volume I [S] (2020) edition Beijing: China Medical Science Press, 2020. [in Chinese]
6. Wang XD (2023) The intellectual origin, conceptual connotation, and contemporary development of medicine and food homology [J]. J Nanjing Univ Chin Med 39(9):809–813 [in Chinese]
7. Shan F, Huang LQ, Guo J et al (2015) A review of the history and development of medicine and food homology [J]. Chin Bull Life Sci 27(8):1061–1069 [in Chinese]
8. Huang LQ, He CN, Ma P et al (2022) Strategic considerations on the industrial development of medicinal and edible substances in China [J]. Strategic Study Chin Acad Eng 24(6):81–87 [in Chinese]
9. Yao RY, He CN, Xiao PG (2023) Food and medicine continuum' in the East and West: Old tradition and current regulation [J]. Chin Herb Med 15(1):6–14
10. Heinrich M, Yao RY, Xiao PG (2022) Food and medicine continuum'—Why we should promote cross-cultural communication between the global East and West [J]. Chin Herb Med 14(1):3–5
11. Xiao XH, Wang JB, Zhao YL et al (2010) The thermodynamics view of Chinese medicine nature and its practice [J]. China J Chin Materia Med 35(16):2207–2213 [in Chinese]
12. Wei GH, Zhang FC, Fu XJ et al (2020) Quantitative modeling of constituent similarity of Chinese medicines and prediction analysis of their cold–hot natures [J]. Data Anal Knowl Discovery 4(5):75–83 [in Chinese]
13. Jiang XB, Jiang NN (2003) The quantification of cold and hot nature of Chinese medicines and the minimum-quantity principle of cold–hot quantification [J]. J Tradit Chin Med 44(7):548–550 [in Chinese]
14. Rosch E (1978) Principles of categorization [M]//Rosch E, Lloyd B B. Cognition and Categorization. Lawrence Erlbaum Associates, Hillsdale, pp 27–48
15. Yang GF Paradigm reframing in medicine–food homology theory: Constructing the dual-track cognition and systemic mapping model [R/OL]. 2025 (manuscript/preprint). [in Chinese]
16. Yang GF (2025) Establishment and validation of a three-dimensional quantitative evaluation system for the nature bias of plant-derived food materials [R/OL]. Preprint. 10.21203/rs.3.rs-8381999/v1[in Chinese]
17. Yang GF (2025) A three-dimensional evaluation of animal-derived food materials and a triple-component formulation synergy system [R/OL]. Preprint. 10.21203/rs.3.rs-7960666/v1[in Chinese]
18. Wen Y, Li HW, Liu FB, Hou ZK (2017) Methodological progress in the quantitative study of Chinese medicine nature [J]. China J Traditional Chin Med Pharm 32(3):1181–1183 [in Chinese]
19. Zhou Y, Xu B (2021) New insights into molecular mechanisms of Cold or Hot nature of food: When East meets West [J]. Food Res Int 145:110442
20. Stevens SS (1946) On the theory of scales of measurement [J]. Science 103(2684):677–680
21. Mokkink LB, Terwee CB, Patrick DL et al (2010) The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments [J]. Qual Life Res 19(4):539–549


## Figures

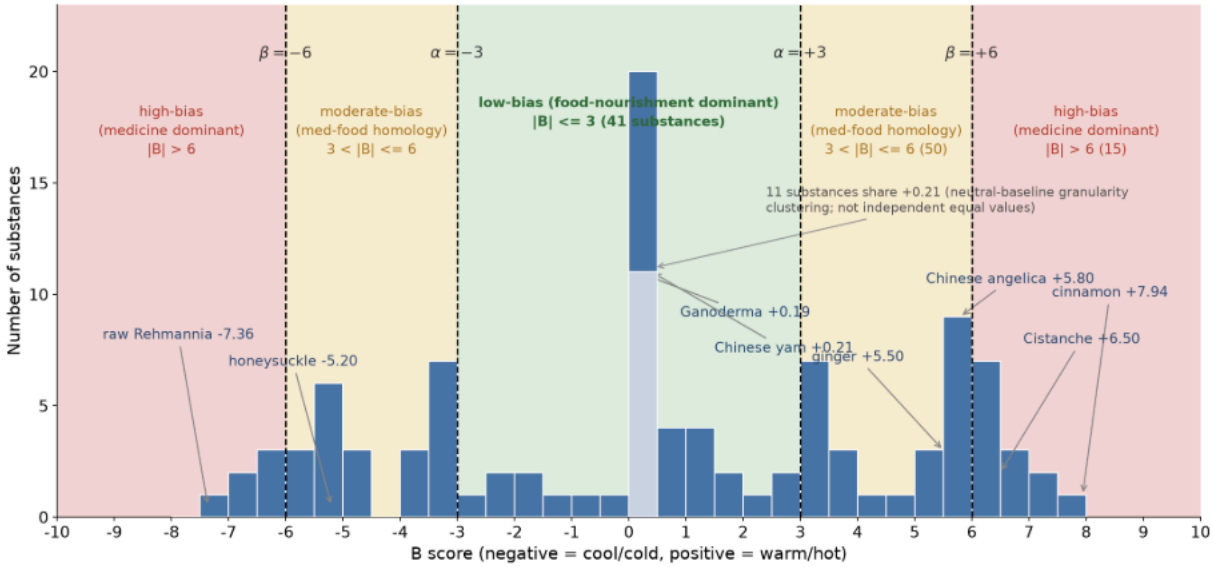



**Figure 1**

*Distribution of the nature bias scores (B) of the 106 medicinal and edible substances and the cut-offs for the three bias zones (the light-coloured bar is the +0.21 neutral-reference granularity cluster, 11 substances, not independent item-by-item measured equal values)*

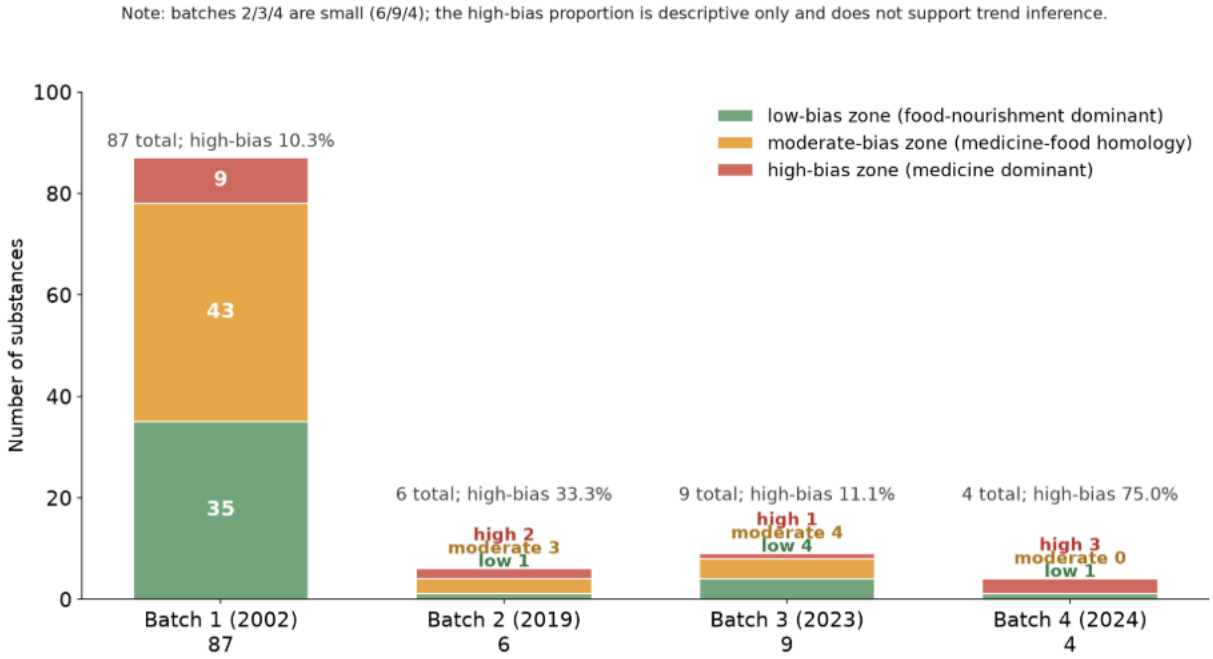



**Figure 2**

*Distribution of the catalogue substances of each batch across the three bias zones (n=106; later batches are small and are shown descriptively only, not for trend inference)*

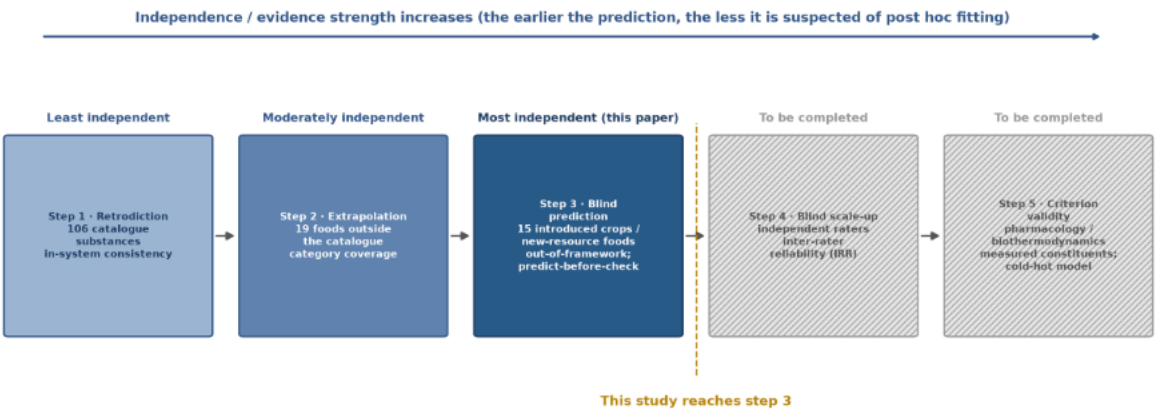



**Figure 3**

*The three-level ladder of predictive validity: retrodiction 106 → extrapolation 19 → blind prediction 15, with independence increasing at each level*