Profiling Self-Reported Metabolic Symptom Patterns Among Urban Delhi NCR Adults:

A Descriptive and Psychometric Study of a Digital Screening Instrument

Deepti Anand[1], and Arun Kumar Chaudhary[2]

[1]Co-Founder & Director, IDCONS TECHNOVA PRIVATE LIMITED; Managing Editor, IDC India Research Journal (IDCIJST)

[2]Chairman, IDC India- Innovation Development Council; Chairman & Co-Founder, GRASA Millets & Foods Pvt. Ltd.

Author Note

Deepti Anand https://orcid.org/0009-0003-7685-466X

Arun Kumar Chaudhary https://orcid.org/0009-0009-0032-4981

The authors are affiliated with IDC India Innovation Development Council and GRASA Millets & Foods Pvt. Ltd., the organisations that operate the screening instrument examined in this study; this is a material conflict of interest. See the Conflict of Interest Statement for full disclosure.

Correspondence concerning this article should be addressed to IDC India- Innovation Development Council, 5th Floor, Wing A, Statesman House, Barakhamba Road, Connaught Place, New Delhi – 110001, India. Email: editorial@idcijst.com

This manuscript is prepared for submission to IDCIJST — Innovation Development Council's International Journal of Science and Technology, ISSN (Online): 3139-7395, published by IDC India Innovation Development Council, New Delhi.

Abstract

Metabolic syndrome, polycystic ovary syndrome (PCOS), and related metabolic dysfunctions have reached high and rising prevalence in urban Indian populations, with several community-based studies from Delhi NCR reporting metabolic syndrome prevalence between roughly 30% and 45% depending on the diagnostic criteria and occupational subgroup studied (Pathak & Aeri, 2026; Sharma et al., 2025; Sinha et al., 2013). Low-cost, symptom-based screening instruments are increasingly used outside clinical settings to triage individuals toward nutrition or lifestyle programmes, but the psychometric performance of such instruments in Indian populations is rarely reported. This study reports self-reported symptom-endorsement rates ,at both the individual-item and aggregated four-domain level ,generated by a 12-item screening instrument administered to an urban Delhi NCR sample, and evaluates the instrument's internal consistency; it does not estimate the population prevalence of metabolic syndrome, PCOS, glycaemic dysregulation, or any other diagnosed condition. This was a cross-sectional, retrospective analysis of self-administered screener responses (N = 94) collected via a WhatsApp-linked digital intake form between April and September 2026. The instrument comprises four three-item domains digestive/gut function, energy and metabolic rate, hormonal and reproductive signals, and glycaemic regulation each scored 0–3 per item (0–9 per domain). Internal consistency was assessed using Cronbach's alpha at the full-scale and subscale level, corrected item-total correlations, and inter-domain Pearson correlations. The most common self-reported dominant profile was Slow Metabolic Rate (31.9%), followed by Hormonal

Metabolic (27.7%), Gut-Metabolic (24.5%), and Glycaemic Dysregulation (16.0%). Full-scale internal consistency was acceptable (Cronbach's α = 0.72, 12 items), but consistency at the individual four-item-domain level was weak to moderate (α range 0.33–0.57), and inter-domain correlations (r = 0.21–0.43) indicated that domains share substantial common variance rather than functioning as fully independent constructs. The instrument's practice of assigning a single named 'dominant profile' should be reconsidered: the observed symptom clustering across domains is consistent with a broader clinical view of metabolic dysfunction as a single interconnected process rather than four separable conditions. A validation study comparing screener scores against biochemical confirmation is the recommended next step.

Keywords: metabolic syndrome; symptom screening; digital health screening; internal consistency; Cronbach's alpha; urban India; Delhi NCR; PCOS; glycaemic dysregulation

Profiling Self-Reported Metabolic Symptom Patterns Among Urban Delhi NCR Adults

India is undergoing a well-documented epidemiological transition in which non-communicable, metabolically driven conditions insulin resistance, type 2 diabetes, polycystic ovary syndrome (PCOS), non-alcoholic fatty liver disease, and functional gut disorders have become common even among apparently healthy, employed, urban adults (Krishnamoorthy et al., 2020). Community and workplace studies from the Delhi National Capital Region (NCR) illustrate the scale of the problem: a study of women in South Delhi found a metabolic syndrome prevalence of 29.6% using NCEP/ATP-III criteria (Sinha et al., 2013); a 2026 cross-sectional study of white-collar employees across Delhi NCR and Bengaluru found metabolic syndrome in 40.5% of participants, with a higher point estimate in Delhi NCR (44.6%) than Bengaluru (36.6%) (Pathak & Aeri, 2026); and a large 2025 survey of college-going women in Delhi NCR reported a high burden of PCOS-consistent symptoms, particularly among urban, higher-education, and nuclear-family subgroups (Sharma et al., 2025).

Two features distinguish these studies from the data examined here. First, they rely on clinical or anthropometric measurement fasting glucose, lipid panels, waist circumference, blood pressure rather than self-report alone. Second, they are purpose-built research studies with ethics review, defined sampling frames, and, in most cases, probability or systematic sampling. A large and growing share of real-world metabolic screening in India, however, happens outside this research infrastructure: through app- and WhatsApp-based intake questionnaires that route users

toward commercial wellness or nutrition programmes on the basis of self-reported symptoms alone, with no laboratory confirmation and no independent validation of the underlying symptom taxonomy.

This raises two distinct questions that the present study addresses. The first is descriptive: how common is self-identification with each of the screener's symptom areas both individual items and aggregated domains , among respondents in this urban Delhi NCR sample? This is a question about self-reported symptom endorsement, not about the prevalence of any diagnosed condition, and clinically confirmed figures from the same geography (cited above) are used only as descriptive backdrop, not as a claim this instrument reproduces them. The second is methodological: does a screening instrument of this kind behave as a coherent multi-domain measure, or does it function, statistically, as a single general symptom-burden scale wearing a four-profile label? This second question matters because the instrument evaluated here assigns each respondent a single 'dominant profile' based on which three-item domain scores highest ,a classification decision that is only meaningful if the four domains are behaving as sufficiently distinct constructs.

This paper does not evaluate, and makes no claims about, the clinical or commercial nutrition-programme recommendations generated downstream of the screener; that layer is explicitly outside the scope of this analysis.

Research Statement

In a self-selected sample of urban Delhi NCR adults screened through a real-world digital intake channel, what proportion self-report symptoms, at both the individual-item level and the aggregated four-domain level, consistent with each of the screener's four metabolic-symptom categories; and does the instrument's four-domain scoring structure hold up statistically as measuring genuinely distinct symptom clusters, or does it primarily reflect one general symptom-burden dimension? This study reports self-reported symptom-endorsement rates only. It does not estimate, and should not be read as estimating, the population prevalence of metabolic syndrome, PCOS, glycaemic dysregulation, or any other diagnosed condition ,no clinical or biochemical confirmation of any respondent's symptoms was obtained, and none is claimed. Where clinically confirmed prevalence figures from the same geography are cited elsewhere in this paper, they are included only as descriptive context for how common these symptom areas are as a public health concern in this setting, not as a benchmark this instrument is shown to reproduce.

Literature Review

Metabolic Syndrome and Its Components in Urban India

Metabolic syndrome ,the clustering of central obesity, dyslipidaemia, hypertension, and impaired glucose regulation ,is estimated to affect roughly one in three Indian adults, with meta-analytic prevalence estimates varying by diagnostic criteria and setting (Krishnamoorthy et al., 2020). Prevalence is consistently higher in urban than rural populations and among

individuals with sedentary occupations. In a community sample of women of lower socioeconomic status in South Delhi, Sinha et al. (2013) reported a prevalence of 29.6% by NCEP/ATP-III criteria and 20.4% by IDF criteria, rising with age and caloric intake. In a more recent, occupationally defined sample, Pathak and Aeri (2026) found a 40.5% prevalence among white-collar private-sector employees aged 30–50 across Delhi NCR and Bengaluru, with elevated fasting blood glucose significantly more common in the Delhi NCR sub-sample, and independent associations with age, self-reported dieting, and poor sleep quality (Pittsburgh Sleep Quality Index > 5).

These figures provide the epidemiological backdrop against which self-reported screening data from the same geography should be read ,but they are not directly comparable to the present dataset, since they rely on measured anthropometric and biochemical parameters rather than symptom self-report.

Hormonal and Reproductive Metabolic Dysfunction: PCOS in Delhi NCR

PCOS has been described as reaching near-epidemic proportions among urban Indian women, attributed in the literature to obesogenic diets, sedentary lifestyle, and the broader metabolic-syndrome cluster (Sharma et al., 2025). In a stratified cross-sectional survey of 1,164 college-going women aged 18–25 in Delhi NCR, Sharma et al. (2025) found that age ≥20 years, higher education, East Indian or immigrant ancestry, and nuclear family structure were each associated with increased PCOS risk, and argued that the region's population density and socioeconomic heterogeneity make it a particularly important setting for this research. This

supports treating the 'Hormonal Metabolic' domain of the present screener as capturing a real and regionally salient symptom cluster, even though the screener itself cannot diagnose PCOS.

Self-Report Symptom Screening: Uses and Limitations

Self-administered, Likert-style symptom screeners are attractive for large-scale, low-cost triage because they require no laboratory infrastructure and can be distributed digitally at near-zero marginal cost, a genuine advantage in a health system where formal diagnostic testing remains inaccessible to much of the population. However, convenience-sampled, self-report instruments administered outside a clinical encounter, and distributed via consumer messaging platforms such as WhatsApp rather than research-grade survey infrastructure, are well understood to carry substantial selection bias, cannot support causal or diagnostic claims, and should be described explicitly as non-diagnostic and population non-representative.

Reliability Assessment of Multi-Item Screening Instruments

When a screening instrument groups items into sub-domains and assigns respondents to categories based on relative sub-domain scores, as this instrument does, it is standard practice to report the internal consistency of each sub-domain, not only of the instrument as a whole. Cronbach's alpha is the most widely used index of internal consistency for this purpose (Tavakol & Dennick, 2011). Tavakol and Dennick (2011) note that alpha is a function of both the average inter-item correlation and the number of items in a scale; short sub-scales (here, three items each)

are structurally prone to lower alpha values even when item content is coherent, which is an important interpretive caveat for the Discussion below. Conventionally, $\alpha \geq 0.70$ is treated as an acceptable threshold for research use, though Tavakol and Dennick caution against treating this threshold mechanically without considering scale length and dimensionality.

Method

Design

This is a retrospective, cross-sectional, descriptive and psychometric analysis of existing screener response data. No new data collection occurred for this study; the dataset was extracted from an operational digital intake system.

Instrument

The instrument is a 12-item, self-administered, English/Hindi bilingual questionnaire (the 'Metabolic Screener') divided into four three-item domains: (A) Digestive Function, (B) Energy & Metabolism, (C) Hormonal & Reproductive Signals, and (D) Blood Sugar & Cravings. Each item offers four response options scored 0 (least symptomatic) to 3 (most symptomatic), so each domain has a possible range of 0–9. Respondents' single highest-scoring domain is labelled their 'dominant profile'; where a second domain scores within a defined margin of the top score, a 'secondary profile' is also recorded. The instrument was designed for operational triage purposes

and has not, to the authors' knowledge, undergone independent psychometric validation prior to this analysis.

Setting and Participants

Respondents accessed the screener through a WhatsApp-linked digital form associated with a Delhi NCR-based nutrition programme. The dataset (N = 94) spans submissions dated 17 April 2026 to 11 September 2026. Geographic and demographic scope (Delhi NCR; adult age range) is asserted at the programme/catchment level; the exported dataset itself contains no individual-level age, sex, or location field, despite the source instrument requesting them, and this is treated as a limitation below rather than a verified sampling frame.

Sampling was non-probability and self-selected: individuals who encountered the screener through the programme's own outreach channels and chose to complete it. No formal inclusion/exclusion criteria, sample-size calculation, or comparison group were applied at the point of data collection, since the instrument was not originally designed as a research tool.

Ethical Considerations

This analysis uses data that were originally collected for operational/commercial screening purposes, not under a research protocol. Given the retrospective, non-interventional, and fully de-identified nature of the analysis, review and approval for secondary academic use was granted internally by Dr. Arun Kumar Chaudhary, MBBS, Chairman, IDC India ,Innovation Development Council, in his capacity as the senior physician and governance authority

responsible for the underlying programme (see the Ethics Statement below). All data used were fully de-identified prior to analysis: no names, phone numbers, or other direct identifiers were used in any computation, and none appear in this manuscript or its tables.

Statistical Analysis

Descriptive statistics (mean, standard deviation, range) were computed for each of the 12 items and each of the four three-item domain totals. Internal consistency was assessed using Cronbach's alpha, computed as:

$$\alpha = (k / (k - 1)) \times [1 - (\Sigma\sigma_i^2 / \sigma\square^2)]$$

where k is the number of items, $\Sigma\sigma_i^2$ is the sum of individual item variances, and $\sigma\square^2$ is the variance of the summed scale score (Tavakol & Dennick, 2011). Alpha was computed separately for each three-item domain and for the full 12-item instrument. Corrected item-total (item-rest) correlations were computed for each item against the sum of the remaining items in its domain, to identify items contributing weakly to their assigned domain. Pearson product-moment correlations were computed between all six pairs of domain totals to assess the degree of shared variance across domains. All analyses were conducted in Python (NumPy); no imputation was performed, as the dataset contained no missing item-level responses.

Results

Ninety-four respondents were included. Table 1 shows the distribution of dominant and secondary profiles; Table 2 shows item-level response distributions; Tables 3–5 report the

instrument's internal-consistency and correlation structure, referenced throughout this section and the Discussion below.

The most common dominant profile was Slow Metabolic Rate (31.9%), followed by Hormonal Metabolic (27.7%), Gut-Metabolic (24.5%), and Glycaemic Dysregulation (16.0%); see Table 1 and Figure 1. Dominant-profile assignment was independently recomputed from raw item responses against the instrument's own scoring rubric, and all 94 assignments matched exactly (0 discrepancies), confirming the scoring engine applies its stated rule consistently. Item-level endorsement rates (Table 2, Figure 2) ranged from 10.6% (Q3, food intolerances) to 37.2% (Q4, post-meal energy, and Q8, mood fluctuation).

Full-scale internal consistency was acceptable (Cronbach's $\alpha = 0.72$, 12 items), but each individual three-item domain showed weak-to-moderate consistency ($\alpha = 0.33–0.57$; Table 3, Figure 3). All six domain pairs were positively correlated ($r = 0.21–0.43$; Table 4, Figure 4), ranging from $r = 0.21$ (Digestive–Glycaemic) to $r = 0.43$ (Energy/Metabolism–Hormonal) , indicating that respondents scoring highly on one domain tend to score at least moderately on the others, i.e., the domains are not orthogonal. Corrected item-total correlations (Table 5) show that Q3 (food intolerances) contributes negligibly to the Digestive domain ($r = 0.06$) and Q12 (fasting glucose/HbA1c/insulin history) contributes weakly to the Glycaemic domain ($r = 0.18$); both are conceptually distinct, lower-frequency items within their domains, which may explain their weaker statistical alignment with domain-mates capturing more common day-to-day symptoms.

Discussion

Profile Distribution in Context (Deepti Anand)

The near-even spread of dominant profiles across the sample (16.0–31.9%, with no single domain dominating overwhelmingly) is worth reading as a substantive finding in its own right, not only as a methodological limitation. Metabolic dysfunction rarely originates in a single organ system: lifestyle, demographic, and temporal factors act at the level of the cell, and disturbances at that level propagate outward across digestive, hormonal, glycaemic, and organ-specific function together, rather than in one isolated channel. This view is increasingly reflected in the clinical literature itself ,PCOS, historically treated as a reproductive-endocrine condition, is now widely described as fundamentally a metabolic disease driven by insulin resistance, present even in lean patients and independent of body weight (Genazzani & Genazzani, 2023), with insulin resistance identified as a unifying mechanism underlying both metabolic and reproductive dysfunction (Prosperi & Chiarelli, 2025). On this understanding, 'metabolic' is better treated as an umbrella term for an interconnected physiological process than as one category standing alongside 'hormonal,' 'digestive,' or 'glycaemic' as though they were independent. A lifestyle-driven disturbance such as insulin resistance would be expected, on this basis, to surface as gut symptoms, hormonal symptoms, and glycaemic symptoms together in the same person, in sequence or in parallel, rather than in only one domain.

Read this way, the positive inter-domain correlations reported in Table 4 ($r = 0.21–0.43$) are not only a psychometric caution about the instrument's four-way classification, discussed next ,they may also be a genuine physiological signal consistent with this systemic view of

metabolic health. The two explanations are not mutually exclusive: the subscales may be both too short to show strong standalone reliability and be picking up a real, shared underlying process. Either way, the practical implication for how this class of instrument should present its output is the same as the one reached independently below from a clinical perspective ,a single named 'dominant profile' understates how interconnected these symptom areas actually are.

What the Reliability Analysis Shows

The full 12-item instrument showed acceptable internal consistency ($\alpha = 0.72$), but each individual three-item domain showed weak-to-moderate consistency ($\alpha = 0.33–0.57$), and all six domain pairs were positively correlated ($r = 0.21–0.43$). Read together with Tavakol and Dennick's (2011) point that alpha is partly a function of scale length, this pattern is consistent with the twelve items substantially tapping a single underlying 'general metabolic-symptom burden' factor, with the four-way 'dominant profile' split representing a finer distinction than the item-level data cleanly supports.

This statistical ambiguity has a direct clinical corollary. As discussed next, assigning a respondent a single named category on the basis of a three-item, weakly-consistent subscale risks conveying more categorical certainty than the underlying data warrant. The practical concern is not merely psychometric: a respondent told they are, specifically, a 'Hormonal Metabolic' or

'Glycaemic Dysregulation' type may reasonably treat that label as a sufficient explanation for their symptoms and stop looking further for an underlying cause ,including causes that fall outside any nutrition programme's scope and require clinical evaluation.

Clinical and Public Health Perspective (Arun Kumar Chaudhary)

From clinical experience, patients presenting with metabolic complaints at a general hospital setting rarely show an isolated, single-system pattern. Digestive complaints, low energy, hormonal irregularity, and blood-sugar-related symptoms typically occur together in the same patient rather than in clean isolation, consistent with the inter-domain correlations reported in Table 4. This is expected physiologically ,metabolic, digestive, and endocrine systems are interdependent, so a symptom cluster spanning more than one 'domain' is the norm rather than the exception in real patients, not a flaw unique to this instrument.

The clinical concern with a tool of this kind is not that it collects self-reported symptoms ,that has genuine value as an awareness and triage aid, particularly at low cost and at scale. The concern is what happens after a respondent receives a single named 'dominant profile.' Given that the four sub-scales here are only weakly-to-moderately internally consistent and are all positively correlated with one another, a label of this kind risks being read by both the respondent and any downstream advisor as more diagnostically specific than it is. In practice, the risk is that a patient treats the label as a sufficient explanation and does not pursue further clinical evaluation ,which matters most for the minority of respondents whose symptoms in fact reflect a diagnosable condition (for example, frank diabetes, thyroid dysfunction, or PCOS) that requires medical

management rather than a nutrition or lifestyle programme alone. Accordingly, an instrument of this kind is best positioned to end-users as a general awareness aid that may prompt a clinical consultation, not as a categorical assessment in itself, and ideally should carry an explicit prompt to seek clinical evaluation wherever multiple domains score in the higher range ,which, per Table 4, appears to be the more common presentation, not the exception.

Public Health Policy and Health-Economics Perspective

Two established frameworks are useful for situating this instrument within public health policy, rather than treating it only as a private commercial tool. The first is the classic Wilson and Jungner (1968) criteria for evaluating any population screening programme, still the most widely cited reference point in screening policy 60 years on (Andermann et al., 2008). Among the ten original criteria are the requirements that a suitable test or examination for the condition should exist, that the test should be acceptable to the population, and ,most directly relevant here ,that the natural course of the condition should be adequately understood and the test's validity established before large-scale use. The reliability findings reported above speak directly to this last criterion: a four-category classification built on subscales with α as low as 0.33 has not yet demonstrated the validity that formal screening-policy frameworks treat as a precondition for population-level use, independent of whether the underlying health conditions it gestures toward (metabolic syndrome, PCOS, glycaemic dysregulation) are themselves well understood.

The second is the health-economics case for low-cost digital screening in a resource-constrained system. There is a real and growing evidence base showing that

mHealth-based tools can be cost-effective or cost-saving relative to conventional clinical screening pathways, particularly for reach into populations that would otherwise not be screened at all (Iribarren et al., 2017). This is a genuine argument in favour of instruments like the one examined here: near-zero marginal distribution cost via WhatsApp is a real advantage in a health system where laboratory-based screening remains inaccessible to much of the population. However, this economic case is a case for well-validated low-cost screening, not a substitute for validation ,the cost-effectiveness literature on mHealth tools generally evaluates instruments already shown to correlate with clinical outcomes, which is precisely the step this instrument has not yet undergone. The policy-relevant conclusion, then, is not that self-report digital screening should be discarded on cost-benefit grounds, but that its low cost is only a genuine public health asset once paired with the kind of reliability evidence this paper has begun to generate.

Limitations

This study has several limitations that should be weighed carefully by readers and reviewers. First, the sample is a non-probability, self-selected convenience sample drawn from a single programme's outreach channel; it is not representative of Delhi NCR or urban India more broadly, and no formal sample-size or power calculation was performed. Second, geographic and demographic scope (Delhi NCR; adult age range; sex distribution) is asserted at the programme/catchment level and could not be verified respondent-by-respondent within the exported dataset, which contained no age, sex, or location fields despite the source instrument requesting them. Third, all data are self-reported over a three-month symptom-recall window,

with no biochemical, anthropometric, or clinical corroboration; the four profile labels ('Gut-Metabolic', 'Slow Metabolic Rate', 'Hormonal Metabolic', 'Glycaemic Dysregulation') should not be read as equivalent to, or predictive of, clinical diagnoses such as metabolic syndrome, PCOS, or pre-diabetes. Fourth, the four three-item subscales showed weak-to-moderate internal consistency ($\alpha = 0.33–0.57$) and were all positively inter-correlated, which ,as discussed above ,raises an open question about whether the instrument's four-profile classification is statistically well supported by the item data. Fifth, the design is cross-sectional with a single administration per respondent; no test–retest reliability could be assessed, and no outcome or follow-up data exist to evaluate predictive or clinical validity. Sixth, as noted above under Ethical Considerations, this analysis uses data originally collected for operational rather than research purposes; formal documentation of that secondary-use approval should be finalised before submission. Finally, the analysis excludes the instrument's downstream nutrition-programme routing logic entirely, by design, and no claims are made in this paper about the clinical or commercial value of that layer.

Conclusion

This study finds that a widely deployable class of digital metabolic screening tools shows acceptable reliability as a single composite measure but weak reliability at the level of its four individual domains, and that these domains are not statistically independent. Metabolic dysfunction rarely respects organ-system boundaries ,a view increasingly reflected in the clinical literature on conditions such as PCOS ,and the tendency of respondents' scores across all four symptom domains to move together, rather than independently, is consistent with this understanding. Read this way, the instrument's difficulty cleanly separating 'hormonal' from 'metabolic' from 'glycaemic' symptoms may be less a flaw unique to this tool and more a faithful reflection of how these systems actually behave together in the body. The practical implication is the same either way: the instrument's current practice of assigning each respondent a single named 'dominant profile' should be reconsidered, since it risks conveying more categorical certainty than either the statistical or the clinical evidence presented here supports.

This study reports self-reported symptom-endorsement rates only, at the individual-item and aggregated domain level, and makes no claim about the prevalence of metabolic syndrome, PCOS, glycaemic dysregulation, or any other diagnosed condition. The natural next step is a validation study comparing this or a similarly structured screener's scores against biochemical confirmation (fasting glucose/HbA1c, hormone panels) in a subsample, which would test directly whether the symptom clustering observed here reflects genuine shared physiology, measurement noise, or both.

Table 1

Distribution of Dominant and Secondary Metabolic-Symptom Profiles (N = 94)

Profile	Dominant, n (%)	Secondary, n (%)
Slow Metabolic Rate	30 (31.9%)	5 (5.3%)
Hormonal Metabolic	26 (27.7%)	13 (13.8%)
Gut-Metabolic	23 (24.5%)	0 (0.0%)
Glycaemic Dysregulation	15 (16.0%)	7 (7.4%)

Note. Dominant-profile assignment was independently recomputed from raw item responses against the instrument's own scoring rubric; all 94 assignments matched exactly (0 discrepancies).

Figure 1

Distribution of Dominant Metabolic-Symptom Profiles (N = 94)

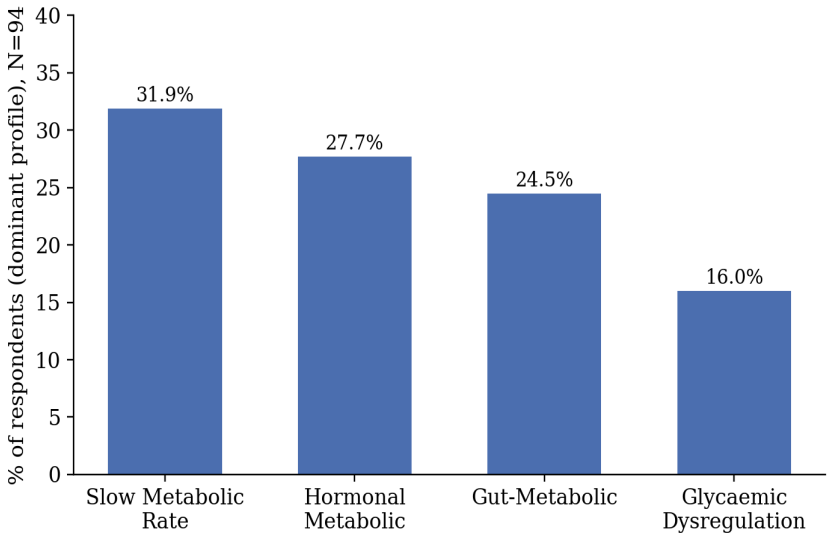


Table 2

Item-Level Response Distribution and Proportion Scoring in the Two Most Symptomatic Response Categories (C/D, Score ≥ 2)

Item	Domain	Mean (SD)	% scoring ≥2
Q1 ,Post-meal digestion	Digestive	0.86 (0.96)	22.3%
Q2 ,Bowel regularity	Digestive	1.00 (0.88)	25.5%
Q3 ,Food intolerances	Digestive	0.45 (0.77)	10.6%
Q4 ,Post-meal energy	Energy/Metabolism	1.33 (1.00)	37.2%
Q5 ,12-month weight change	Energy/Metabolism	0.76 (0.84)	17.0%
Q6 ,Morning energy	Energy/Metabolism	0.88 (0.96)	23.4%
Q7 ,Cycle/monthly consistency	Hormonal	0.86 (0.84)	18.1%
Q8 ,Mood fluctuation	Hormonal	1.31 (0.95)	37.2%
Q9 ,Skin/hair changes	Hormonal	1.03 (0.87)	28.7%
Q10 ,Sugar/carb cravings	Glycaemic	1.22 (0.95)	30.9%
Q11 ,Delayed-meal symptoms	Glycaemic	1.09 (0.92)	27.7%
Q12 ,Fasting glucose/HbA1c/insulin history	Glycaemic	0.39 (0.81)	11.7%

Figure 2

Item-Level Symptom Endorsement by Domain (N = 94)

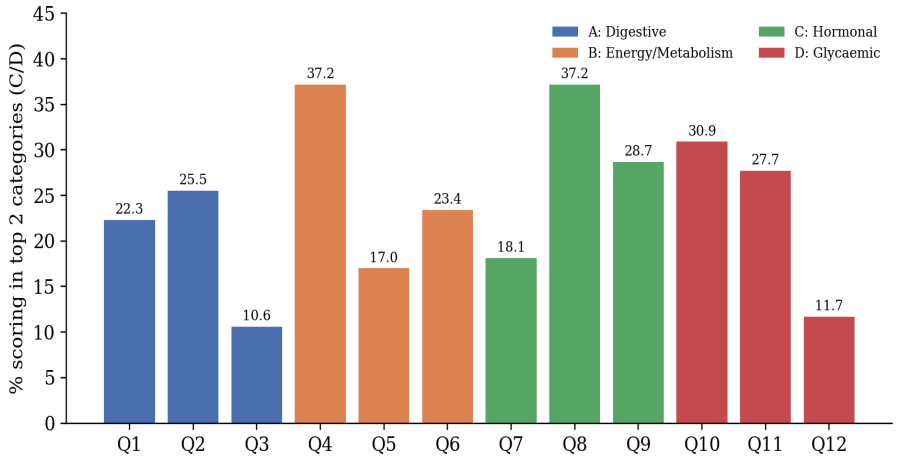


Table 3

Domain-Level Descriptive Statistics and Internal Consistency

Domain (items)	Mean (SD)	Range	Cronbach's α
A: Digestive (Q1–3)	2.31 (1.71)	0–7	0.33
B: Energy/Metabolism (Q4–6)	2.97 (2.02)	0–9	0.54
C: Hormonal (Q7–9)	3.20 (1.95)	0–8	0.57
D: Glycaemic (Q10–12)	2.70 (1.78)	0–8	0.36
Full instrument (12 items)	—	0–29	0.72

Figure 3

Internal Consistency (Cronbach's α) by Domain and Full Instrument, Against the Conventional α ≥ 0.70 Threshold

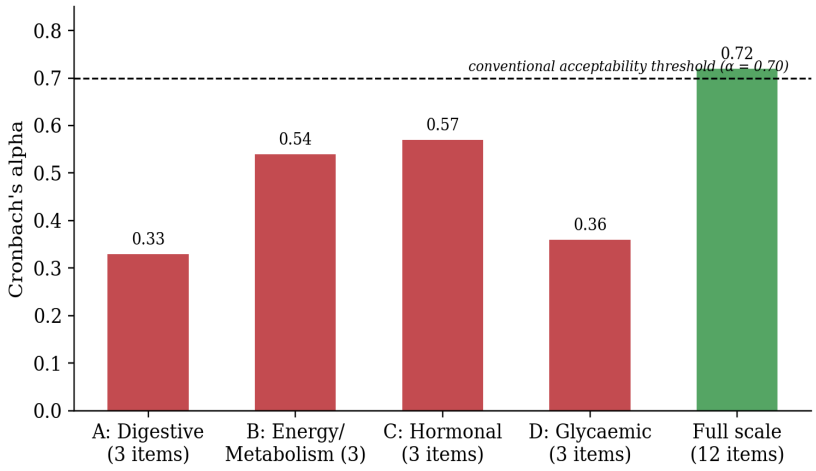


Table 4

Inter-Domain Pearson Correlations

	A	B	C	D
A: Digestive	—	0.40	0.31	0.21
B: Energy/Metabolism	0.40	—	0.43	0.34
C: Hormonal	0.31	0.43	—	0.32
D: Glycaemic	0.21	0.34	0.32	—

Note. All correlations were positive, ranging from r = 0.21 (Digestive–Glycaemic) to r = 0.43 (Energy/Metabolism–Hormonal), indicating the domains are not orthogonal.

Figure 4

Inter-Domain Correlation Matrix (Pearson r)

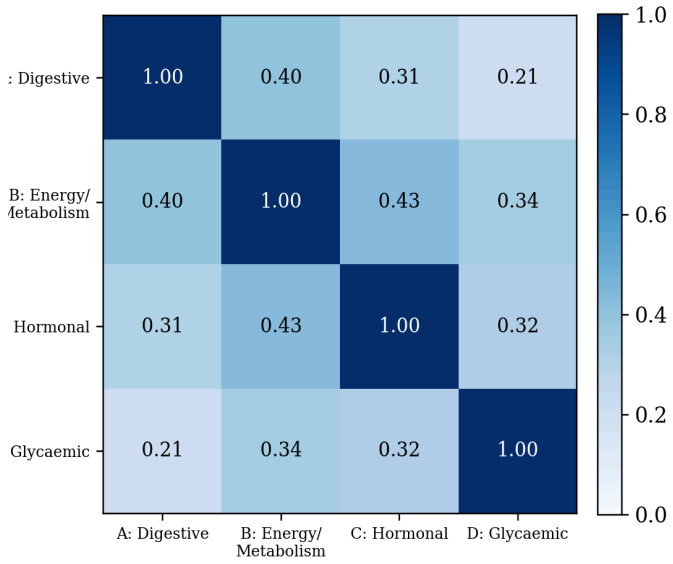


Table 5

Corrected Item-Total (Item-Rest) Correlations Within Domain

Domain	Item 1	Item 2	Item 3
A: Digestive	0.28 (Q1)	0.24 (Q2)	0.06 (Q3)
B: Energy/Metabolism	0.38 (Q4)	0.24 (Q5)	0.45 (Q6)
C: Hormonal	0.43 (Q7)	0.25 (Q8)	0.48 (Q9)
D: Glycaemic	0.25 (Q10)	0.20 (Q11)	0.18 (Q12)

Note. Q3 (food intolerances) contributes negligibly to the Digestive domain (r = 0.06) and Q12 (fasting glucose/HbA1c/insulin history) contributes weakly to the Glycaemic domain (r = 0.18).

Author Contributions

Conceptualization, Methodology, Data curation, Writing – original draft, Writing – review & editing ,Deepti Anand. Clinical and public health interpretation and ethics approval for secondary data use ,Arun Kumar Chaudhary. Formal analysis (statistical computation prepared with AI assistance, reviewed by both authors).

Conflict of Interest Statement

The authors are affiliated with the organisations that operate the screening instrument and associated nutrition programme from which this dataset originates (IDC India and GRASA Millets & Foods Pvt. Ltd.). This is a material conflict of interest, disclosed here in full, and readers should weigh the findings ,particularly the favourable framing of the instrument's overall (full-scale) reliability ,in light of this affiliation.

Ethics Statement

This study used retrospective, fully de-identified data originally collected for programmatic screening purposes. Given the retrospective, non-interventional, and de-identified nature of the analysis, review and approval for secondary academic use was granted internally by

METABOLIC SCREENING PROFILES

Dr. Arun Kumar Chaudhary, MBBS, Chairman, IDC India ,Innovation Development Council, in his capacity as the senior physician and governance authority responsible for the underlying programme.

References

Andermann, A., Blancquaert, I., Beauchamp, S., & Déry, V. (2008). Revisiting Wilson and Jungner in the genomic age: A review of screening criteria over the past 40 years. Bulletin of the World Health Organization, 86(4), 317–319. https://doi.org/10.2471/BLT.07.050112

Genazzani, A. D., & Genazzani, A. R. (2023). Polycystic ovary syndrome as metabolic disease: New insights on insulin resistance. touchREVIEWS in Endocrinology, 19(1), 71–77. https://doi.org/10.17925/EE.2023.19.1.71

Iribarren, S. J., Cato, K., Falzon, L., & Stone, P. W. (2017). What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. PLOS ONE, 12(2), e0170581. https://doi.org/10.1371/journal.pone.0170581

Krishnamoorthy, Y., Rajaa, S., Murali, S., Rehman, T., Sahoo, J., & Kar, S. S. (2020). Prevalence of metabolic syndrome among adult population in India: A systematic review and meta-analysis. PLOS ONE, 15(10), e0240971. https://doi.org/10.1371/journal.pone.0240971

Pathak, V., & Aeri, B. T. (2026). Prevalence of cardiometabolic risk factors and metabolic syndrome in white-collar private sector employees from two Indian metropolitan cities. Discover Public Health, 23, Article 830. https://doi.org/10.1186/s12982-026-02195-4

Prosperi, S., & Chiarelli, F. (2025). Insulin resistance, metabolic syndrome and polycystic ovaries: An intriguing conundrum. Frontiers in Endocrinology, 16, Article 1669716. https://doi.org/10.3389/fendo.2025.1669716

Sharma, A., Sarwal, Y., Devi, N. K., et al. (2025). Polycystic Ovary Syndrome prevalence and associated sociodemographic risk factors: A study among young adults in Delhi NCR, India. Reproductive Health, 22, Article 61. https://doi.org/10.1186/s12978-025-02019-9

Sinha, S., Misra, P., Kant, S., Krishnan, A., Nongkynrih, B., & Vikram, N. K. (2013). Prevalence of metabolic syndrome and its selected determinants among urban adult women in South Delhi, India. Postgraduate Medical Journal, 89(1048), 68–72. https://pubmed.ncbi.nlm.nih.gov/23112218/

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. International Journal of Medical Education, 2, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Wilson, J. M. G., & Jungner, G. (1968). Principles and practice of screening for disease (Public Health Papers No. 34). World Health Organization. https://iris.who.int/handle/10665/37650