Am J Transl Res 2021;13(11):12557-12574
www.ajtr.org /ISSN:1943-8141/AJTR0118543

*Original Article*

# Modeling factors critical for implementation of precision medicine at health systems-level: an IRT approach

John JO Mogaka, Moses J Chimbari

*Department of Public Health Medicine, University of KwaZulu-Natal, Durban, South Africa*

Received July 20, 2020; Accepted December 13, 2020; Epub November 15, 2021; Published November 30, 2021

**Abstract:** Background: Through recent advances in omics technologies, precision medicine (PM) promises to fundamentally change the way we approach health, disease and illness. Imperative applications of omics-based biomarkers are gradually moving from research to clinical settings, with huge long-term clinical and public health implications. Whereas much of research in PM is mainly focused on basic biomedical discoveries, currently there is little research on the clinical implementation of omics biomarkers, especially at health systems level. Aim and Methods: This study investigated the application of multidimensional item response theory (IRT) models to validate a hypothesized PM implementation measurement model. This is a contribution to PM implementation at health systems level. Data obtained through an item-sort procedure involving 496 observations from 124 study participants formed the basis of a 22-item PMI measurement model. Conclusion: Statistical significance of the bifactor model suggests PM implementation may have to be examined using factors that reflect a single common underlying implementation construct, as well as factors that reflect unique variances for the identified four content-specific factors.

**Keywords:** Precision medicine, omics technologies, implementation model, IRT, bifactor modelling

## Introduction

Through medically actionable information gleaned from biological data, recent groundbreaking achievements in genomics are poised to fundamentally shift the biopsychosocial approach to health and disease as we know it today. Using tools collectively referred to as omics technologies [1], genomics, epigenomics and other complete sets of molecular parameters offer deeper insights into the functions and interactions within the human cell (genome) and its environment. Such molecular depictions offer crucial insights into biological processes underlying various human disease and health statuses. Furthermore, the simultaneous analysis of omics data through whole genome sequencing (WGS) and gene expressions studies [2] have been used to explain the relationships among multiple variables in disease etiology, particularly in assessing genotype-phenotype relationships [3]. **Figure 1** illustrates five main omics technologies and their corresponding biomarkers.

Whereas OBM discovery has received much attention through basic biomedical research [4], less attention has been placed on their translation through the biomarker translation pipeline for routine clinical application, especially at health systems level. This discrepancy is more pronounced in resource-poor countries [5, 6]. Considering the long-term implications of PM, interrelated context- and innovation-level barriers and/or facilitation factors for its implementation must be considered for its implementation in resource-poor settings to fully realize its benefits. Evidence in existing implementation literature [7-9] show that moving most health innovations from point of discovery to point of application is not only a challenging process, but also multifaceted, with no known implementation guidelines [10]. Implementing PM through OBM, therefore, not only involves attention to the biomarkers' analytical and clinical validity, but also a wide array of multi-level contextual factors [10].

Although there is an array of implementation frameworks that could potentially inform PM implementation [11-14], their psychometric measures need sufficient validation if the models were to be of adequate use in quantitative investigation of PM implementation [15]. This

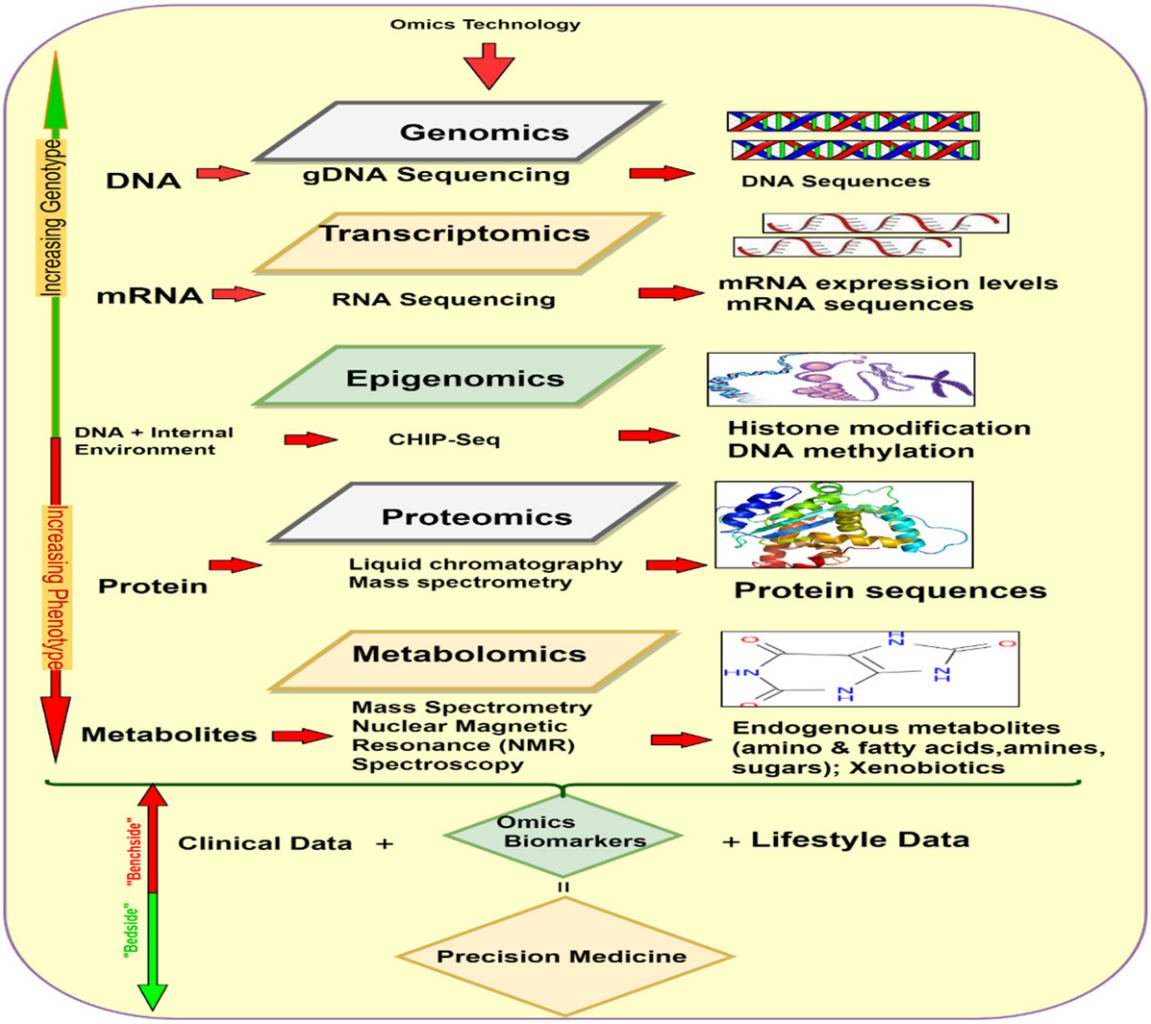


**Figure 1.** Five of the main omics technologies and corresponding biomarkers: Genomics: DNA sequences using genomic DNA (gDNA-sequencing); Transcriptomics: messenger RNA (mRNA) sequencing and mRNA expression level using RNA-sequencing; Epigenomics: DNA methylation and histone modification using ChIP-sequencing; Proteomics: protein sequence using liquid chromatography and mass spectrometry; Metabolomics: metabolites using mass spectrometry and NRM.

study sought to fill this measurement gap in PM implementation, especially in resource-poor settings. The study builds upon converging evidence and concepts contained in four related implementation frameworks to guide its selection of factors that influence the integration of OBMs into clinical and population health settings. The implementation frameworks leveraged upon are the Promoting Action on Research Implementation in Health Services (PARIHS) [14, 16-18]; Consolidated Framework for Implementation Research (CFIR) [11, 19, 20]; Practical, Robust Implementation and Sustainability Model (PRISM) [12]; and Reach, Effectiveness, Adoption, Implementation and Maintenance (RE-AIM) [13]. The study constructed latent variables (constructs) that were hypothesized to have varying influences on PM implementation at different levels of analysis. Their relationship with corresponding observable indicators was then statistically tested. The overarching question answered in this paper, however, is: What is the internal structure or dimensionality of the PMI measurement tool? The answer informs whether the PMI tool that was developed is a general measure of implementation of PM or it is a measure of several context-specific implementation factors at the health systems level. Essentially, this is a measure of scale dimensionality that underlies

the PMI scale and the degree to which ignoring the pertinent dimensionality (uni- or multi-dimensionality) degrades (if at all) resultant dimensional solutions.

A range of psychometric techniques exists for determining whether a set of observable indicators of a latent factor is sufficiently unidimensional or multidimensional. This paper used item response theory (IRT) [21] implemented in R (version 3.6.1) statistical environment to model the dimensionality of the PM implementation (PMI) construct. IRT is a general framework for specifying the functional relationship between a respondent's underlying latent trait level (equivalent to 'factor score' in factor analysis).

*The hypothesized measurement model*

Although various models (further explained in the methods section) were used to analyze the dimensionality pattern of the PMI measurement model, most were unsuitable in explaining variables that may simultaneously tap multiple dimensions of the PMI measurement model, since such models only concern one dimension of analysis at a time. However, the bifactor model [22, 23] was found to be useful, especially in examining distortions in the unidimensional model. The utility of the general factor and subscales was robust in defining, justifying, and clarifying relationships that may fail to hold across factors, even though they may have done so at particular levels of analysis. The bifactor model, also called nested factors model [24], more accurately captured the PMI construct and was therefore adopted.

As explicated in the following sections, through existing literature various factors that collectively influence PM implementation at health systems level were identified. Secondly, we leveraged known implementation science tools to group these factors into four categories: characteristics of the omics-based biomarkers (specifically their clinical validity and utility); organizational support through resource allocation for PM implementation, (financial, technical and human skill); public genomic knowledge and engagement reflected by user response; and genomics-friendly health and biomedical regulatory frameworks. The various implementation research tools relied upon for the conceptualization of the four factors included Consolidated Framework for Implementation Research (CFIR) [11], Promoting Action on Research Implementation in Health Services (PARIHS) [14] and Reach, Effectiveness, Adoption, Implementation and Maintenance (RE-AIM) framework [13, 25]. These four latent variables were therefore hypothesized and abstracted to form the basis of the health-systems-level PM implementation model.

*Characteristics of omics biomarkers*

Omics technologies contribute to precision medicine (PM) through the identification of appropriate omics-based biomarkers (OBMs) [26]. OBMs, in turn, are useful in various biomedical applications such as neonatal screening, disease diagnostics, disease risk prediction, patient stratification and pharmacodynamic or treatment response monitoring [27].

The possibility of explaining genotype-phenotype relationships at the molecular level has resulted in an increased interest in the use of OBMs to improve health outcomes [28]. For instance, in addition to the use of conventional biomarkers (e.g. histopathological assessment of tumors), novel omics-based biomarkers e.g. BRCA1 mutation (also known as 'breast cancer 1, early onset') is required to optimally benefit from newly discovered cancer therapeutics [29]. Epidermal growth factor receptor (EGFR), human epidermal growth factor receptor 2 (Her2/neu) and anaplastic lymphoma kinase (ALK) are other excellent examples of target markers whose upstream changes through the genome, transcriptome, epigenome and proteome levels are often associated with varying patient-specific clinical responses to therapies [30]. The use of OBMs in clinical and population health has variously been referred to as genomic medicine [31], personalized medicine [32], precision medicine [33] and precision health [34]. Although these terms are usually used interchangeably, precision medicine (PM) refers to the coupling of patient-specific omics data with their lifestyle and environmental factors to offer predictive, individualized and preventive medicine [35-37].

In literature, a genetic test generally refers to an assay of a single gene for a single indication, a gene-panel test refers to multi-gene tests for a common indication, whole exome sequencing (WES) assays the entire exome, while whole genome sequencing (WGS) assesses the entire genome. Although these differentiations exist,

for simplicity and pragmatic reasons, this study referred to the term 'omics biomarker' (OBM) as either a single gene biomarker such as HBF in sickle cell anemia [38, 39], or a multiple biomarker score indicating a multivariate omics model or a multi-gene combined score such as Oncotype DX or MammaPrint [40, 41]. This is because the ultimate idea behind these biomarkers is the generation of data that constitutes a single explanatory variable.

Whereas characteristics of omics biomarkers are intrinsically objective and easily observable, this study focused on their perceived advantages. Despite apparent advantages associated with omics biomarkers, varying perceptions about their effectiveness remain, largely due to less-conclusive clinical validity. This includes questions around the biomarkers' reliability, interpretation of test results and quality standards [42]. Based on existing literature, six items were selected as indicators for 'Omics Biomarkers' (OBM) construct as illustrated in **Table 1**.

*Organizational support*

For PM implementation to succeed, organizational and technical resources must be expended. Biomedical research mostly deals with sensitive genetic information, requiring a set of methodological and data governance complexities that require time, budget, and resources to implement. Clinical research organizations commit resources to collect and store biomaterials through biobanking. In addition, meaningful public engagement and data collection must be budgeted for. Organizational support, therefore, is a factor that can influence PM implementation through OBMs. Based on existing literature, six items were selected as indicators for organizational support construct as illustrated in **Table 1**.

*Public genomic awareness*

Although use of biomarkers is beneficial in circumstances such as identifying susceptibility for inherited conditions, this often implies genetic profiling, which can lead to anticipation, skepticism and concern at the personal level [43]. For example, individuals found to have an inherited susceptibility to cancer after undertaking a genetic test or tests on their relatives, might face social discrimination or stigmatization. In addition, family members may be disenfranchised by the very process of genetic testing, particularly if some wish to undertake the test and others do not, or if some individuals found out information about their own risk through genetic test results of other family members, or if those with normal test results experienced survivor guilt. Genetic test results could also be misinterpreted [44] or create a desperation leading to the use of unproven medical therapies [45]. Besides, many aspects of testing-nature of test, mode of inheritance of a condition, person tested, social or medical context-might contribute to their acceptability or rejection among involved individuals, both as providers and as patients.

Therefore, the level and quality of public genomic awareness/engagement give rise to differing genomic experiences, attitudes, perceptions and values among members of the public, both as individuals and health providers. This also influences PMI outcomes. Six items were selected to be indicative of the construct of 'Public Genomic Awareness' as seen in **Table 1**.

*Implementation outcomes*

Indicators for the implementation outcome construct were deduced from the RE-AIM Framework [13]. Using elements of RE-AIM, those in research or practice can make use of necessary information to justify adoption of the biomarker, and how to maintain it if adopted or widen its reach (penetration) into a given service setting. The hypothesized implementation model explains what influences PM implementation outcomes. Such evidence can inform the design and execution of implementation strategies that aim to change relevant determinants. Four items were selected as measures of the construct 'Implementation Outcomes' (**Table 1**).

The four latent variables, their level of analysis and observed indicators (measures) are given in **Table 1**.

## Methods

*Study participants*

The snowball sampling method, a non-probability sampling method, was applied in identifying potential participants from population of interest for this study. The seed (initial) sample population composed of individuals affiliated to

**Table 1.** Constructs, level of analysis and selected indicators

| Construct/level of analysis | No of Items |
|---|---|
| Characteristics of "Omics biomarkers" Innovation level construct | X1. The omics test has been used among people similar to the present target population |
| | X2. Genetic counselling is part of the procedures in administering the omics test |
| | X3. It is easy to obtain the right amount of sample needed for the omics test from participants/patients |
| | X4. It is easy to obtain the right quality sample needed for the omics test from participants/patients |
| | X5. The time it takes to obtain results for intended use after the omics test is reasonable |
| | X6. There are clear instructions on how to obtain samples for the omics test |
| "Organizational Support" Organization level construct | X7. The amount of money dedicated by our organization for implementing this omics test is enough |
| | X8. The amount of physical space dedicated by our organization for the purposes of implementing this omics test is enough |
| | X9. As the volume, variety and data availability associated with the implementation of this omics test grows, the organization increases the capacity of its data systems to facilitate analysis and user needs |
| | X10. To match latest data handling needs in the implementation of this omics test, the organization regularly hires people with relevant skills |
| | X11. There exist clear procedures on how to store, access, manage and share available data associated with this omics test in the organization |
| | X12. The organization has capacity to handle (e.g. capture, validate., store, and process) data associated with this genetic/omics biomarker |
| "Public Genomic Awareness/ Acceptance": i.e. Acceptability and compatibility of the omics tests among Involved Individuals Individual level construct | X13. Getting buy-in from the public (patients, providers) in carrying out this biomarker testing is easy |
| | X14. Participants easily give consent to take samples for biomarker testing |
| | X15. Publicity and free information publicly available about the genetic/omics biomarker cause potential users to willingly ask or look for it |
| | X16. Using this genetic/omics test has been regarded by practitioners as an appropriate mechanism for patient management (e.g. aid in drug dosage decisions, in carrying fetuses to term or carry out prophylactic surgery) |
| | x17. There is a considerable 'pushback' from practitioners as they feel the genetic/omics test is not consistent with their skills, role, or job expectations |
| | x18. Target individuals feel that the genetic/omics test is in line with their family members' wishes, desires and expectations |
| "Implementation outcomes" System level construct | X19. The genetic/omics test is yet to be used as a routine practice within its intended service setting |
| | X20. Practitioners are more willing to order the genetic/omics test more often whenever deemed necessary |
| | X21. The number of eligible persons accessing the genetic/omics test is far less than the total number potentially in need of the service |
| | X22. So far, the authorities that are supposed to acquire the biomarker testing service have communicated a decision to fully fund its roll out |

various academic institutions and organizations known to be involved in molecular/genetic testing and omics-based biomarkers across Africa. Participant recruitment was also based on their involvement in the field of precision medicine and/or biomarker-related biomedical research as indicated by contribution to the field through publications in peer-reviewed journals and/or attendance of related academic conferences. Therefore, some participants were identified in precision or genomic medicine-related academic conferences and invited to participate. Guided by general principles of the Nuremberg Code, the Declaration of Helsinki and institutional review board permit obtained from the University of KwaZulu-Natal (BREC Permit Ref No BE513/18), the study package was distributed via email to potential participants between June and July 2019. Since this study contained negligible risk of potential embarrassment or other ethical dilemmas that are usually associated with snowball sampling in many other studies, initial participants were encouraged to forward the email containing the study package to their colleagues. The study package included an invitation letter with study description, consent form and a link to an online questionnaire. Even though there were a total of 124 study participants who were recruited for the study, each participant made a total of four observations. Therefore, for compactness and statistical purposes, the data was formatted to have one observation per row (i.e., for 124 participants, there were 124*4 = 496 observations). The eligibility criteria aimed at identifying participants who may fairly represent the demographic characteristics of the main study population on whom the measurement tool was to be used.

Snowball sampling was deemed the ideal sampling method for the study as it was expected that initial study participants would likely know others in the same industry or academic circles as themselves and hence, could collegially inform others about the study and its potential benefits. However, despite this advantage, use of this method meant that it was not possible to determine the sampling error based on the obtained sample. We remedied this apparent weakness in sample size determination in two ways. First, we made *a priori* assumptions about various clusters of the population of interest-those in precision medicine related research or practice. For example, investigators involved in clinical genomics were grouped under “academics” cluster while commercial direct-to-consumer (DTC) genetic test providers and hospital-based pharmacogenomics testing providers were grouped under “industry”. Snow-ball sampling was continued until there was population cluster saturation [46]. Secondly, although there is no definitive number for a pilot study sample size according to Hunt, Sparkman, & Wilcox [47], to increase statistical power and degrees of freedom for our study, we applied a Monte Carlo simulation (MCS) approach as described in Mooney [48]. MCS studies are computer-driven experimental investigations in which certain parameters, such as population means and standard deviations that are known a priori, are used to generate random (but plausible) sample data [48]. This method of generating and analyzing data treats the collected research sample as a “population reservoir” from which a large number of random samples are drawn with continuous replacement such that the probability of selection for any given case remains equal over every random draw [49]. We requested 2,000 iterations, drawn with replacement from the original data set of 124 cases (our empirical sample).

*Demographic characteristics of participants*

There was a total of 124 study participants in the study. There were 12 more female participants than males (10%). Those with research qualifications (PhDs and equivalent) were the majority (55%), while slightly a quarter of the participants had medical qualifications. The rest of participants’ qualities varied greatly and are summarized in **Figure 2**.

*Procedure*

A card-sort task approach [50] was used to obtain comparative item-rating data. Respondents were presented with a set of cards that contained constructs clearly defined using everyday language and a set of categories onto which the concepts on the cards were to be mapped, as explained in Angleitner, John, & Lohr [51]. The respondents were asked to read each item and assign it to the construct or concept it best indicated, in their judgment. Participants were required to rank the cards according to their own understanding and preference, indicating their perception on concepts presented on each card. If participants deemed it necessary to make any changes as to where

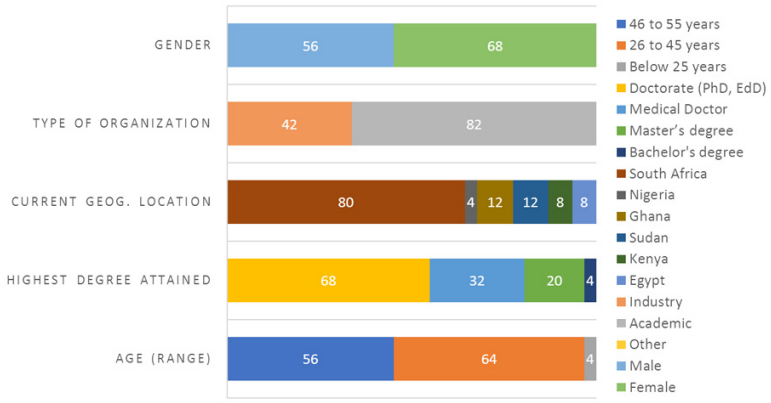


**Figure 2.** Demographic characteristics of study participants.

a card previously assigned should be reassigned to, they could do so. Although simple in practice, this exercise generated enough data that was used to assess the desired person characteristics, construct validity and scale reliability. The actual sorting of the items was, however, performed on an online platform hosted and supported by Optimal Workshop (https://www.optimalworkshop.com/optimalsort). Besides offering convenience to study participants, the online platform enhanced data security and confidentiality. A total of 22 items were to be assigned into 4 categories initially provided for, with participants allowed to create own separate construct categories if they so desired. The results about number of items per category and the items' ranking order (positioning) within their respective categories were then extracted and analyzed. Data was subjected to item response theory (IRT) analysis.

*The data*

In summary, the data are an output of an exercise where respondents were required to place items arranged as cards into specific categories and ranked in order of perceived importance. The cards related to specific attributes indicative of wider factors thought to influence PM implementation at health systems level. This meant that if an item on the card was regarded as the most important in a given category, it was ranked 1st, and the next ranked 2nd and so on. If a card was not placed in a particular category, it was scored zero (0) in that category. The data did not contain any missing values because all the fields of the online item sort tool were set as "required" to eliminate non-responses.

*Statistical considerations*

We took a model comparison approach within the IRT framework to identify the most sensible model to fit to the 22-item PMI scale. Four models were constructed. The first involved a unidimensional graded response model where all item-responses reflected the same common general latent trait (Model A). In other words, Model A consisted of all the 22 PMI items as explained by a single latent factor, PMI. The second model considered a multidimensional graded response model, consisting of more than one correlated PMI dimensions (correlated traits, Model B). Essentially, Model B allowed subsets of the 22 items to be explained by one of the four domain factors of PMI as shown in **Table 1**. The latent factors were then allowed to covary with one another. The third model involved a higher-order model where dimensions (first-order traits) were correlated with a higher-order trait (second-order trait, i.e., PMI), Model C. Finally, the fourth model specified a bifactor model, involving a general (primary) trait that explains all the items while simultaneously taking into account specific (sub-scale) traits, Model D. Model D therefore consisted of all items that were simultaneously explained by a general latent trait "g" while a subset of items were explained by domain specific traits (i.e., OBM, OrG, UsR and ImO). The 22 items were grouped to reflect specific content parcels, as described in Tellegen & Waller [52]. Analysis to assess the model's dimensionality was carried out as explained in Reise, Moore & Maydeu-Olivares [53]. Each model was fit to the data using Multi-dimensional Item Response Theory, MIRT [54] package in R version 3.6.1. The package was chosen based on its suitability in the analysis of dichotomous and polytomous response data using unidimensional and multidimensional latent trait models under the Item Response Theory paradigm. It was ideal for our study for confirmatory bi-factor and two-tier analyses as well as multiple group analysis for modeling item and person covariates. The Marginal Maximum Likelihood (MML) [55] methods were used to estimate the IRT models. However, for those that could not properly converge when using MML (i.e., Model B and C), Metropolis-

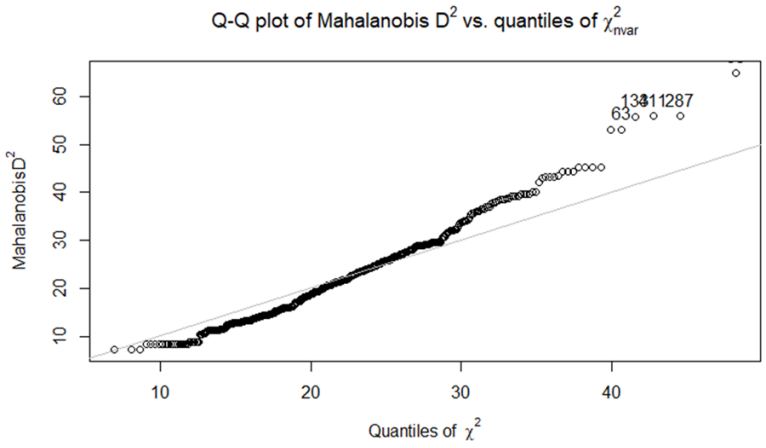


**Figure 3.** Quantile-Quantile (Q-Q) plot describing distribution of data and Mahabolis distances.

Hastings Robbins-Monro (MHRM) [56] was used. Conditional independence was done according to Chen and Thissen's (1997) [57] standardized local dependency (LD) $\chi^2$ statistic.

Model fit to the data was evaluated based on four indices: the root mean square error of approximation (RMSEA); Standardized Root Mean Square Residual (SRMSR); the Tucker-Lewis index (TLI) and the comparative fit index (CFI) [58].

## Results

*Correlations and descriptive statistics*

Descriptive statistics of the data were obtained, including an analysis of outliers. As shown in **Figure 2**, six observations were detected as outliers. However, these outliers did not significantly alter results and so they were included in the dataset. **Figure 3** is a quantile-quantile (Q-Q) plot formed by the data and shows the distribution of the data against the expected normal distribution. Since the observations did tend to approximate a straight diagonal line with minimal deviations, we concluded that our data has a general normal multivariate distribution. This normal distribution pattern justified use of the marginal maximum likelihood estimation method (MML).

We used the R package "corrplot" version 0.84 [59] to visualize the data correlation matrix. The package was used because it is good in graphically displaying correlation matrices in finer details, including choosing color to indicate various parameters, text labels and layout. The correlogram in **Figure 4** displays variables in the correlation matrix and how they relate with each other. In the upper triangle, positive correlations are displayed in blue and negative correlations in red color. Color intensity and the size of the circle are proportional to the correlation coefficients, helping to identify "groups" of variables that share a strong relationship with each other (hierarchical clustering). The lower triangular correlation matrix displays the actual correlation values.

*Model dimensionality analysis*

Local item independence tests were performed to ascertain data dimensionality and aid in identifying the most appropriate model for the hypothesized PMI measurement model. **Figure 5** summarizes the unidimensional (Model A), multidimensional correlated (Model C) and bifactor (Model D) models' large positive (+) LD (> 10) and negative (-) LD (< -10) values. **Appendix 1** presents further summaries of all model LD statistics results. **Figure 5** (and **Appendix 1**) indicates that the unidimensional model (Model A) has more positive LDs compared to the other models. Large positive LDs imply unmodeled covariation among some item pairs, suggesting Model A may not have properly specified the number of latent variables that explain the item response patterns, i.e., it did not appropriately account for all covariations. This further implies that item response covariations may be better modeled by a multidimensional model such as the multidimensional (correlated or second order) or the bifactor model. A closer look at **Figure 5** shows that the bifactor Model D is best suited to model the dimensionality of the PMI scale, with a marked improvement in modeling the item response covariations.

*Evaluation of item-level model-data fit, global model-data fit and model comparison*

Item-level model-data fit using the Orlando-Thissen-Bjorner item fit S-$\chi^2$ statistic [60] and bias-corrected using Benjamini-Hochberg (B-H) procedure [61] were used to indicate how well

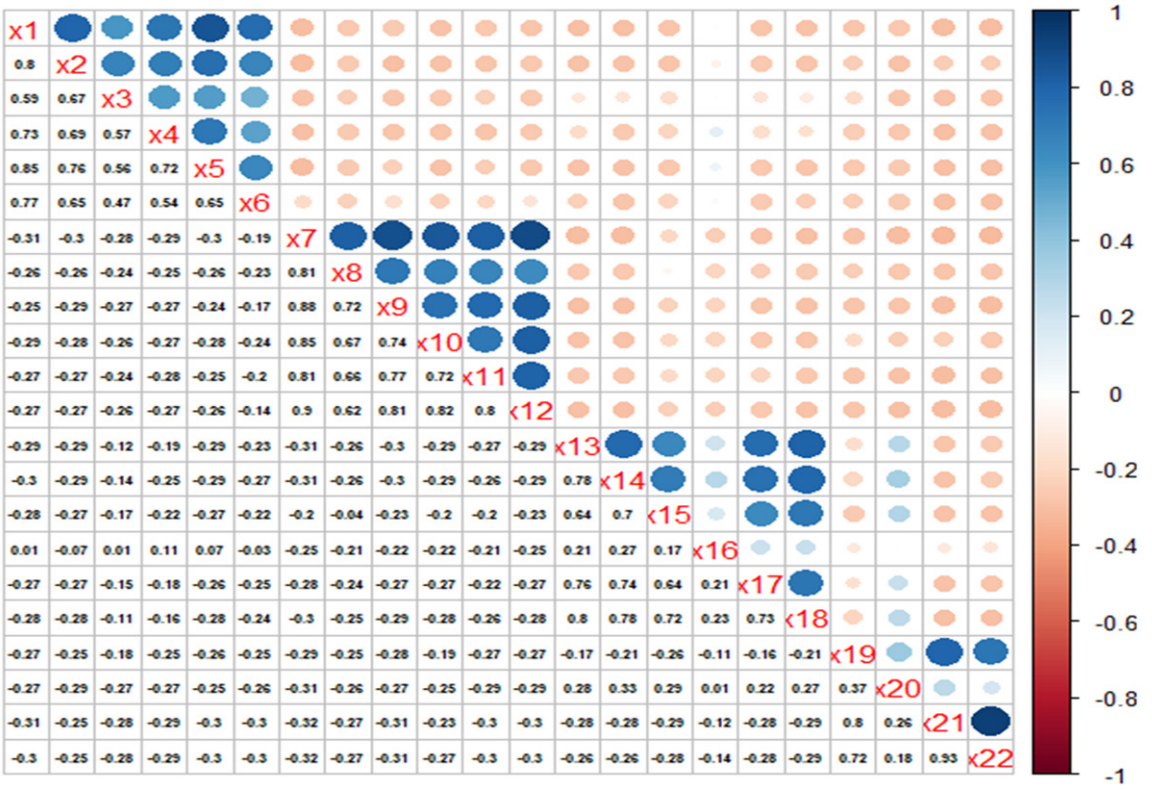


**Figure 4.** Variable correlation matrix.

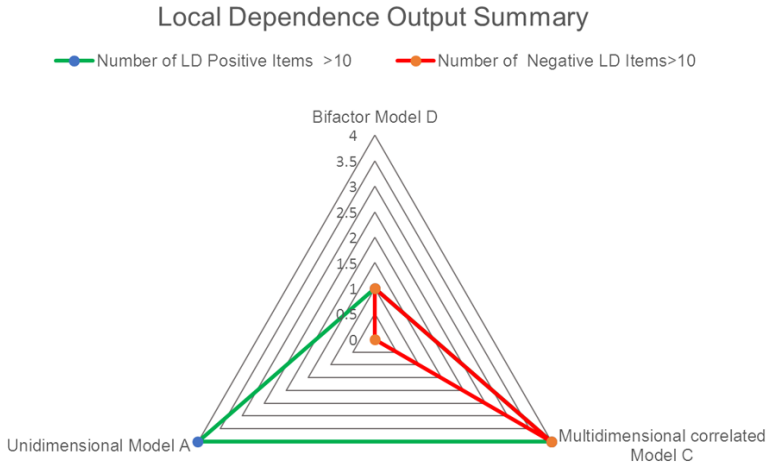


**Figure 5.** LD values for three IRT models.

each of the four models predicted response behavior at the item level (Detailed results are shown in **Appendix 2**). The results showed that all the models except unidimensional showed non-significant values. However, the unidimensional model showed only one item as slightly significant, essentially suggesting that all four IRT models adequately fit the items. However, once again the unidimensional model was discounted as a possible explanation for the data fit since item fit statistics are only accurate under the assumption that the model dimensionality is correctly specified by the IRT model. This follows evidence as presented in **Figure 4**, where, based on LDs, it was shown that the unidimensional model showed more unmodeled item response covariation.

After ascertaining item-level model-data fit, the nested Models A through D were tested and compared in terms of their global goodness of fit statistics. The deviance statistic, given by the difference in the -2 log likelihood (-2LL) between two nested models, was also used. **Table 2** displays a summary of the model comparison and fit results for the four competing IRT models. All methods (BIC, AIC, M2) for determining the best-fitting model agree that the bifactor and multidimensional correlated models appear to be the two most

**Table 2.** Global model-data fit indices and factor loadings

| Model | Bifactor | Multidimensional 2nd order | Multidimensional correlated | Unidimensional |
|---|---|---|---|---|
| M2 | 205.56 | 484.4353 | 168.2866 | 3126.0443 |
| df | 56 | 78.0000 | 72.0000 | 78.0000 |
| p | 0 | 0.0000 | 0.0000 | 0.0000 |
| RMSEA | 0.06277539 | 0.1026 | 0.0622 | 0.2810 |
| SRMSR | 0.21753126 | 0.2167 | 0.2148 | 0.3100 |
| TLI | 0.96918246 | 0.9399 | 0.9846 | 0.5491 |
| CFI | 0.98274218 | 0.9531 | 0.9889 | 0.6483 |
| Log-likelihood | -8013.57 | -8085.72 | -7861.06 | -9253.02 |
| 2LL | -16027.13 | -16171.44 | -15722.12 | -18506.04 |
| Estimated parameters | 197 | 175 | 181 | 175 |
| AIC | 16421.13 | 16521.44 | 16084.12 | 18856.04 |
| BIC | 17249.83 | 17257.59 | 16845.51 | 19592.19 |

Bifactor Model Factor Loadings

| | g | OBM | OrG | PGA | ImO | h2 |
|---|---|---|---|---|---|---|
| x1 | 0.5657 | 0.808 | 0.000 | 0.000 | 0.0000 | 0.973 |
| x2 | 0.3769 | 0.882 | 0.000 | 0.000 | 0.0000 | 0.921 |
| x3 | 0.2032 | 0.880 | 0.000 | 0.000 | 0.0000 | 0.816 |
| x4 | 0.4370 | 0.797 | 0.000 | 0.000 | 0.0000 | 0.826 |
| x5 | 0.5103 | 0.827 | 0.000 | 0.000 | 0.0000 | 0.944 |
| x6 | 0.5425 | 0.727 | 0.000 | 0.000 | 0.0000 | 0.823 |
| x7 | 0.2863 | 0.000 | 0.937 | 0.000 | 0.0000 | 0.960 |
| x8 | 0.1752 | 0.000 | 0.940 | 0.000 | 0.0000 | 0.913 |
| x9 | 0.3224 | 0.000 | 0.899 | 0.000 | 0.0000 | 0.912 |
| x10 | 0.1264 | 0.000 | 0.958 | 0.000 | 0.0000 | 0.934 |
| x11 | 0.2969 | 0.000 | 0.884 | 0.000 | 0.0000 | 0.870 |
| x12 | 0.3281 | 0.000 | 0.911 | 0.000 | 0.0000 | 0.937 |
| x13 | -0.1129 | 0.000 | 0.000 | 0.954 | 0.0000 | 0.923 |
| x14 | -0.1099 | 0.000 | 0.000 | 0.951 | 0.0000 | 0.917 |
| x15 | 0.0810 | 0.000 | 0.000 | 0.931 | 0.0000 | 0.873 |
| x16 | 0.1946 | 0.000 | 0.000 | 0.495 | 0.0000 | 0.283 |
| x17 | 0.0288 | 0.000 | 0.000 | 0.970 | 0.0000 | 0.941 |
| x18 | 0.0319 | 0.000 | 0.000 | 0.980 | 0.0000 | 0.960 |
| x19 | -0.7402 | 0.000 | 0.000 | 0.000 | 0.5163 | 0.814 |
| x20 | -0.5052 | 0.000 | 0.000 | 0.000 | -0.0179 | 0.256 |
| x21 | -0.6364 | 0.000 | 0.000 | 0.000 | 0.7515 | 0.970 |
| x22 | -0.6416 | 0.000 | 0.000 | 0.000 | 0.7240 | 0.936 |

Multidimensional 2nd order Model Factor Loadings

| | OBM | OrG | PGA | ImO | PMI | h2 |
|---|---|---|---|---|---|---|
| x1 | 0.982 | 0.000 | 0.000 | 0.000 | 0 | 0.965 |
| x2 | 0.950 | 0.000 | 0.000 | 0.000 | 0 | 0.902 |
| x3 | 0.854 | 0.000 | 0.000 | 0.000 | 0 | 0.730 |
| x4 | 0.910 | 0.000 | 0.000 | 0.000 | 0 | 0.827 |
| x5 | 0.974 | 0.000 | 0.000 | 0.000 | 0 | 0.949 |
| x6 | 0.903 | 0.000 | 0.000 | 0.000 | 0 | 0.815 |
| x7 | 0.000 | 0.983 | 0.000 | 0.000 | 0 | 0.966 |
| x8 | 0.000 | 0.956 | 0.000 | 0.000 | 0 | 0.914 |
| x9 | 0.000 | 0.958 | 0.000 | 0.000 | 0 | 0.918 |
| x10 | 0.000 | 0.959 | 0.000 | 0.000 | 0 | 0.919 |
| x11 | 0.000 | 0.936 | 0.000 | 0.000 | 0 | 0.875 |
| x12 | 0.000 | 0.966 | 0.000 | 0.000 | 0 | 0.933 |
| x13 | 0.000 | 0.000 | 0.955 | 0.000 | 0 | 0.912 |
| x14 | 0.000 | 0.000 | 0.952 | 0.000 | 0 | 0.907 |
| x15 | 0.000 | 0.000 | 0.929 | 0.000 | 0 | 0.863 |
| x16 | 0.000 | 0.000 | 0.482 | 0.000 | 0 | 0.232 |
| x17 | 0.000 | 0.000 | 0.970 | 0.000 | 0 | 0.940 |
| x18 | 0.000 | 0.000 | 0.980 | 0.000 | 0 | 0.960 |
| x19 | 0.000 | 0.000 | 0.000 | 0.896 | 0 | 0.803 |
| x20 | 0.000 | 0.000 | 0.000 | 0.389 | 0 | 0.151 |
| x21 | 0.000 | 0.000 | 0.000 | 0.985 | 0 | 0.970 |
| x22 | 0.000 | 0.000 | 0.000 | 0.973 | 0 | 0.948 |

Multidimensional Correlated Factor Loadings

| | OBM | OrG | PGA | ImO | h2 |
|---|---|---|---|---|---|
| x1 | 0.977 | 0.000 | 0.000 | 0.000 | 0.954 |
| x2 | 0.935 | 0.000 | 0.000 | 0.000 | 0.873 |
| x3 | 0.824 | 0.000 | 0.000 | 0.000 | 0.679 |
| x4 | 0.889 | 0.000 | 0.000 | 0.000 | 0.790 |
| x5 | 0.964 | 0.000 | 0.000 | 0.000 | 0.930 |
| x6 | 0.880 | 0.000 | 0.000 | 0.000 | 0.774 |
| x7 | 0.000 | 0.972 | 0.000 | 0.000 | 0.945 |
| x8 | 0.000 | 0.936 | 0.000 | 0.000 | 0.875 |
| x9 | 0.000 | 0.936 | 0.000 | 0.000 | 0.877 |
| x10 | 0.000 | 0.938 | 0.000 | 0.000 | 0.879 |
| x11 | 0.000 | 0.908 | 0.000 | 0.000 | 0.824 |
| x12 | 0.000 | 0.948 | 0.000 | 0.000 | 0.899 |
| x13 | 0.000 | 0.000 | 0.943 | 0.000 | 0.889 |
| x14 | 0.000 | 0.000 | 0.940 | 0.000 | 0.883 |

Unidimensional Model Factor Loadings

| | PMI | h2 |
|---|---|---|
| x1 | -0.0498 | 0.00248 |
| x2 | -0.0387 | 0.00150 |
| x3 | 0.0957 | 0.00916 |
| x4 | 0.0574 | 0.00330 |
| x5 | -0.0530 | 0.00281 |
| x6 | -0.0904 | 0.00817 |
| x7 | -0.9800 | 0.96044 |
| x8 | -0.9488 | 0.90017 |
| x9 | -0.9521 | 0.90645 |
| x10 | -0.9516 | 0.90554 |
| x11 | -0.8955 | 0.80195 |
| x12 | -0.9618 | 0.92498 |
| x13 | 0.9608 | 0.92322 |
| x14 | 0.9592 | 0.92010 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| x15 | 0.000 | 0.000 | 0.908 | 0.000 | 0.825 | x15 | 0.8650 | 0.74827 |
| x16 | 0.000 | 0.000 | 0.448 | 0.000 | 0.201 | x16 | 0.5034 | 0.25344 |
| x17 | 0.000 | 0.000 | 0.960 | 0.000 | 0.922 | x17 | 0.9703 | 0.94147 |
| x18 | 0.000 | 0.000 | 0.973 | 0.000 | 0.947 | x18 | 0.9795 | 0.95937 |
| x19 | 0.000 | 0.000 | 0.000 | 0.865 | 0.749 | x19 | 0.1211 | 0.01466 |
| x20 | 0.000 | 0.000 | 0.000 | 0.391 | 0.153 | x20 | 0.5831 | 0.33995 |
| x21 | 0.000 | 0.000 | 0.000 | 0.982 | 0.964 | x21 | 0.0858 | 0.00736 |
| x22 | 0.000 | 0.000 | 0.000 | 0.963 | 0.928 | x22 | 0.0894 | 0.00799 |

Note: M2 = goodness-of-fit statistic; df = degrees of freedom; -2LL = -2 log likelihood/deviance statistic; BIC = Bayesian information criterion; AIC = Akaike information criterion; P = *p* value; RMSEA = root mean square error of approximation; SRMSR = Standardized Root Mean Square Residual; TLI = Tucker-Lewis index; CFI = comparative fit index, OBM = Biomarker characteristics factor; PMI = precision medicine implementation; PGA = public genomic awareness; ImO = Implementation Outcomes; g = general factor; h2 = item communalities.

competing models. However, to find out which of the two most competing models to apply to the hypothesized PMI scale, we utilized prior calculated item-level diagnostics statistics (i.e., LD indices and item-fit statistics) to identify any degrees of misfit. Therefore, based on the global model-data fit (TLI and CFI), model comparison results (AIC and BIC) and LD indices (**Figure 4**), the bifactor model appears the most plausible explanation for the responses in the data. Besides, the theoretical foundation for the PMI model as explained under the 'hypothesized PMI measurement model' section seems congruent with this model choice.

## Discussion

The strong correlation within the "OBM" subscale is in agreement with findings elsewhere about the importance of validating the evidence base of PM [62]. This study's findings are also corroborated by similar works that have expressed concerns especially around how valid and reliable available genetic tests and how well they predict desired outcomes [63, 64]; and what the benefits and harms associated with the clinical use of these tests are [65].

On the other hand, this study has succeeded in confirming the measurement model in terms of its IRT specifications. Most importantly in the present context, the conceptually important question of PMI scale's dimensionality was analyzed to answer the question: does the PMI data set have a strong enough single common factor for IRT, or is it a more complex multidimensional IRT representation? Although it was possible to discredit a unidimensional model for the scale, the study further investigated which of the applicable multidimensional models were most appropriate for the kind of data structure analyzed. The bifactor model was found to be the most plausible model to explain the data structure at hand. This was also supported by factor loadings for the four models as indicated in **Table 2**. However, although factor loading is adequate, the bifactor does not suggest a perfect unidimensional construct. On the other hand, strong content-specific correlations on their own, as suggested by the multidimensional correlated model, do not point to the common PMI construct that the scale was meant to measure. Therefore, the combined PMI bifactor model presents a better and plausible model which can explain the PMI scale as conceptualized. This is because the bifactor model shows how all items simultaneously measure both the common PMI latent trait, and at the same time account for the variance of each item as influenced by groupings of items that tap different aspects of PMI, given the common latent trait. This agrees with the conceptualization of a broader PMI general factor suitable for scaling of individuals on a common underlying continuum. This is also in line with the knowledge that factors that may influence PM implementation at health systems level are likely multifactorial. To aid in efficient decision making related to PM implementation therefore, the PMI measurement model necessarily must account for measures of substantively complex constructs, i.e., measures with diverse content indicators. This implies that the selected bifactor model accounts for content heterogeneity in content indicators, as well as 'narrow' content homogeneity. In the PMI instrument, the measure composed in all the 22 items would be conceptually broader than if we selected out a subset of six items that asked specifically about, say, PGA (i.e., items x4 through x7).

An important question to consider in the conceptualization of the PM implementation mo-

del and subsequent validation is whether any omitted and/or unmeasured factors might be a basis for inferential bias. Intuitively, the superlative solution to this problem would be to reliably account for all relevant factors. Although measuring for such additional variables, building them into the model and statistically controlling for them would potentially be an important strategy for dealing with any associated confounds, it has been shown that such statistical control has a shortcoming in that it is useful only in ruling out specific, known and measurable confounds, rather than an entire class of alternative models. Consequently, and in line with existing literature [66, 67], we considered that in our case, measuring all potentially impactful variables for the model would be impossible to achieve. Instead we considered the operative question of the degree to which any omitted variable potentially biased our interpretation of model results and provided a basis for alternative explanations of our findings. However, since the hypothesized PM implementation model was theoretically underpinned based on existing literature, as well as the use of a research design and empirical data analysis that is suited to account for multi-item scale validation, alternative models to the one presented herein was appropriately ruled out. This is because the multiple indicators used to measure the latent variables for athe model allowed for the modelling of any correlated errors. This was confirmed in the extent to which variables were correlated in **Table 2**.

## Conclusion

This study set out to investigate the dimensionality of the hypothesized PMI measurement model. It explored the use of a bifactor model solution to investigate the multifaceted nature of constructs that measured precision medicine implementation at health systems level. Four subscales that had been hypothesized for the PMI were confirmed in the data analysis using the IRT approach. Each subscale fittingly tapped different facets of the same underlying PMI construct. However, the model solution did not support an outright unidimensional scale. Even though the multidimensional correlated model with idiosyncratic factor loadings that suggested perfect content-specific grouping was competitive, it was discounted based on the theoretical underpinnings of the main construct the PMI scale was meant to measure. The bifactor solution solved this challenge by simultaneously reflecting what is common among the items in their four content-specific parcels and representing individual differences on a single target dimension. The findings therefore support a statistical and practical conclusion pointing to the suitability of the bifactor model, that is the PMI scale reflects a co-occurring PMI general factor and four unique content-specific factors.

## Acknowledgements

We acknowledge College of Health sciences, University of KwaZulu-Natal for general moral and financial support.

## Disclosure of conflict of interest

None.

**Address correspondence to:** John JO Mogaka, Department of Public Health Medicine, College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa. Tel: +27833561374; E-mail: johnmogaka2@gmail.com

## References

[1] Horgan RP and Kenny LC. 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics. Obste Gynaecol 2011; 13: 189-195.

[2] Galperin MY and Koonin EV. From complete genome sequence to 'complete' understanding? Trends Biotechnol 2010; 28: 398-406.

[3] Vernon ST, Hansen T, Kott KA, Yang JY, O'Sullivan JF and Figtree GA. Utilizing state-of-the-art "omics" technology and bioinformatics to identify new biological mechanisms and biomarkers for coronary artery disease. Microcirculation 2019; 26: e12488.

[4] Ptolemy AS and Rifai N. What is a biomarker? Research investments and lack of clinical integration necessitate a review of biomarker terminology and validation schema. Scand J Clin Lab Invest Suppl 2010; 70: 6-14.

[5] Dalvie S, Koen N, McGregor N, O'Connell K, Warnich L, Ramesar R, Nievergelt CM and Stein DJ. Toward a global roadmap for precision medicine in psychiatry: challenges and opportunities. Omics 2016; 20: 557-564.

[6] Dandara C, Huzair F, Borda-Rodriguez A, Chirikure S, Okpechi I, Warnich L and Masimirembwa C. H3Africa and the African life sciences ecosystem: building sustainable innovation. OMICS 2014; 18: 733-739.

[7] Brownson RC, Colditz GA and Proctor EK. Dissemination and implementation research in health: translating science to practice. Oxford University Press 2018.
[8] Glasgow RE, Vinson C, Chambers D, Khoury MJ, Kaplan RM and Hunter C. National Institutes of Health approaches to dissemination and implementation science: current and future directions. Am J Public Health 2012; 102: 1274-1281.
[9] Proctor EK, Landsverk J, Aarons G, Chambers D, Glisson C and Mittman B. Implementation research in mental health services: an emerging science with conceptual, methodological, and training challenges. Adm Policy Ment Health 2009; 36: 24-34.
[10] Chaudoir SR, Dugan AG and Barr CH. Measuring factors affecting implementation of health innovations: a systematic review of structural, organizational, provider, patient, and innovation level measures. Implement Sci 2013; 8: 22.
[11] Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA and Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009; 4: 50.
[12] Feldstein AC and Glasgow RE. A practical, robust implementation and sustainability model (PRISM) for integrating research findings into practice. Jt Comm J Qual Patient Saf 2008; 34: 228-243.
[13] Glasgow RE, Vogt TM and Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999; 89: 1322-1327.
[14] Kitson AL, Rycroft-Malone J, Harvey G, McCormack B, Seers K and Titchen A. Evaluating the successful implementation of evidence into practice using the PARiHS framework: theoretical and practical challenges. Implement Sci 2008; 3: 1.
[15] Tabak RG, Khoong EC, Chambers DA and Brownson RC. Bridging research and practice: models for dissemination and implementation research. Am J Prev Med 2012; 43: 337-350.
[16] Helfrich CD, Damschroder LJ, Hagedorn HJ, Daggett GS, Sahay A, Ritchie M, Damush T, Guihan M, Ullrich PM and Stetler CB. A critical synthesis of literature on the promoting action on research implementation in health services (PARIHS) framework. Implement Sci 2010; 5: 82.
[17] Kitson A, Harvey G and McCormack B. Enabling the implementation of evidence based practice: a conceptual framework. Qual Health Care 1998; 7: 149-158.
[18] Rycroft-Malone J. The PARIHS framework-a framework for guiding the implementation of evidence-based practice. J Nurs Care Qual 2004; 19: 297-304.
[19] Lash SJ, Timko C, Curran GM, McKay JR and Burden JL. Implementation of evidence-based substance use disorder continuing care interventions. Psychol Addict Behav 2011; 25: 238.
[20] Sorensen JL and Kosten T. Developing the tools of implementation science in substance use disorders treatment: applications of the consolidated framework for implementation research. Psychol Addict Behav 2011; 25: 262.
[21] Glockner-Rist A and Hoijtink H. The best of both worlds: factor analysis of dichotomous data using item response theory and structural equation modeling. Structural Equation Modeling 2003; 10: 544-565.
[22] Gibbons RD, Bock RD, Hedeker D, Weiss DJ, Segawa E, Bhaumik DK, Kupfer DJ, Frank E, Grochocinski VJ and Stover A. Full-information item bifactor analysis of graded response data. Multivariate Behav Res 2007; 31: 4-19.
[23] Gibbons RD and Hedeker DR. Full-information item bi-factor analysis. Psychometrika 1992; 57: 423-436.
[24] Gustafsson JE and Balke G. General and specific abilities as predictors of school achievement. Multivariate Behav Res 1993; 28: 407-434.
[25] Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, Ory MG and Estabrooks PA. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. Front Public Health 2019; 7: 64.
[26] Quezada H, Guzmán-Ortiz AL, Díaz-Sánchez H, Valle-Rios R and Aguirre-Hernández J. Omics-based biomarkers: current status and potential use in the clinic. Bol Med Hosp Infant Mex 2017; 74: 219-226.
[27] Majewski IJ and Bernards R. Taming the dragon: genomic biomarkers to individualize the treatment of cancer. Nat Med 2011; 17: 304.
[28] Ghosh D and Poisson LM. "Omics" data and levels of evidence for biomarker discovery. Genomics 2009; 93: 13-16.
[29] Ford D, Easton D, Stratton M, Narod S, Goldgar D, Devilee P, Bishop D, Weber B, Lenoir G and Chang-Claude J. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. Am J Hum Genet 1998; 62: 676-689.
[30] Deyati A, Younesi E, Hofmann-Apitius M and Novac N. Challenges and opportunities for oncology biomarker discovery. Drug Discov Today 2013; 18: 614-624.

[31] Wonkam A and Mayosi BM. Genomic medicine in Africa: promise, problems and prospects. Genome Med 2014; 6: 11.
[32] Harvey A, Brand A, Holgate ST, Kristiansen LV, Lehrach H, Palotie A and Prainsack B. The future of technologies for personalised medicine. N Biotechnol 2012; 29: 625-633.
[33] Ginsburg GS and Phillips KA. Precision medicine: from science to value. Health Aff (Millwood) 2018; 37: 694-701.
[34] Amba V, de Souza MCMR, Kenner C and Borges CM. Precision health contributions to public health: an integrative review. J Public Health Epidemiol 2018; 10: 225-232.
[35] Collins FS and Varmus H. A new initiative on precision medicine. N Engl J Med 2015; 372: 793-795.
[36] Zhang X. Precision medicine. Personalized medicine, omics and big data: concepts and relationships. J Pharmacogenomics Pharmacoproteomics 2015; 6: e144.
[37] Council NR. Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease. National Academies Press 2011.
[38] Damanhouri GA, Jarullah J, Marouf S, Hindawi S, Mushtaq G and Kamal MA. Clinical biomarkers in sickle cell disease. Saudi J Biol Sci 2015; 22: 24-31.
[39] Steinberg MH and Adewoye AH. Modifier genes and sickle cell anemia. Curr Opin Hematol 2006; 13: 131-136.
[40] Xin L, Liu YH, Martin TA and Jiang WG. The era of multigene panels comes? The clinical utility of oncotype DX and mammaprint. World J Oncol 2017; 8: 34.
[41] Grant KA, Myburgh EJ, Murray E, Pienaar FM, Kidd M, Wright CA and Kotze MJ. Reclassification of early stage breast cancer into treatment groups by combining the use of immunohistochemistry and microarray analysis. S Afr J Sci 2019; 115: 1-6.
[42] Dancey JE, Dobbin KK, Groshen S, Jessup JM, Hruszkewycz AH, Koehler M, Parchment R, Ratain MJ, Shankar LK and Stadler WM. Guidelines for the development and incorporation of biomarker studies in early clinical trials of novel agents. Clin Cancer Res 2010; 16: 1745-1755.
[43] Mai Y, Koromila T, Sagia A, Cooper DN, Vlachopoulos G, Lagoumintzis G, Kollia P, Poulas K, Stathakopoulos V and Patrinos GP. A critical view of the general public's awareness and physicians' opinion of the trends and potential pitfalls of genetic testing in Greece. Per Med 2011; 8: 551-561.
[44] Giardiello FM, Brensinger JD, Petersen GM, Luce MC, Hylind LM, Bacon JA, Booker SV, Parker RD and Hamilton SR. The use and interpretation of commercial APC gene testing for familial adenomatous polyposis. N Engl J Med 1997; 336: 823-827.
[45] Burke W, Pinsky LE and Press NA. Categorizing genetic tests to identify their ethical, legal, and social implications. Am J Med Genet 2001; 106: 233-240.
[46] van Rijnsoever FJ. (I can't get no) saturation: a simulation and guidelines for sample sizes in qualitative research. PLoS One 2017; 12: e0181689.
[47] Hunt SD, Sparkman RD Jr and Wilcox JB. The pretest in survey research: issues and preliminary findings. Journal of Marketing Research 1982; 19: 269-273.
[48] Mooney CZ. Monte carlo simulation. Sage Publications: 1997.
[49] Shrout PE and Bolger N. Mediation in experimental and nonexperimental studies: new procedures and recommendations. Psychol Methods 2002; 7: 422.
[50] Anderson JC and Gerbing DW. Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. J Appl Psychol 1991; 76: 732.
[51] Angleitner A, John OP and Löhr FJ. It's what you ask and how you ask it: an itemmetric analysis of personality questionnaires. Personality assessment via questionnaires. Springer; 1986. pp. 61-108.
[52] Tellegen A and Waller NG. Exploring personality through test construction: development of the multidimensional personality questionnaire. The SAGE Handbook of Personality Theory and Assessment: Personality Measurement and Testing (Volume 2) 2008; 2: 261.
[53] Reise S, Moore T and Maydeu-Olivares A. Target rotations and assessing the impact of model violations on the parameters of unidimensional item response theory models. Educ Psychol Meas 2011; 71: 684-711.
[54] Chalmers RP. mirt: a multidimensional item response theory package for the R environment. J Stat Softw 2012; 48: 1-29.
[55] Bock RD and Aitkin M. Marginal maximum likelihood estimation of item parameters: application of an EM algorithm. Psychometrika 1981; 46: 443-459.
[56] Cai L. Metropolis-Hastings Robbins-Monro algorithm for confirmatory item factor analysis. J Educ Behav Stat 2010; 35: 307-335.
[57] Chen WH and Thissen D. Local dependence indexes for item pairs using item response theory. J Educ Behav Stat 1997; 22: 265-289.
[58] Hu Lt and Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999; 6: 1-55.

[59] Wei T, Simko V, Levy M, Xie Y, Jin Y and Zemla J. Package 'corrplot'. Statistician 2017; 56: 316-324.
[60] Bjorner J, Smith K, Stone C and Sun X. IRTFIT: a macro for item fit and local dependence tests under IRT models. Lincoln, RI: Quality Metric Incorporated; 2007.
[61] Benjamini Y and Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 1995; 57: 289-300.
[62] Ersek JL, Black LJ, Thompson MA and Kim ES. Implementing precision medicine programs and clinical trials in the community-based oncology practice: barriers and best practices. Am Soc Clin Oncol Educ Book 2018; 38: 188-196.
[63] Buschmann D, Haberberger A, Kirchner B, Spornraft M, Riedmaier I, Schelling G and Pfaffl MW. Toward reliable biomarker signatures in the age of liquid biopsies-how to standardize the small RNA-Seq workflow. Nucleic Acids Rese 2016; 44: 5995-6018.
[64] Kim-Cohen J, Caspi A, Taylor A, Williams B, Newcombe R, Craig IW and Moffitt TE. MAOA, maltreatment, and gene-environment interaction predicting children's mental health: new evidence and a meta-analysis. Mol Psychiatry 2006; 11: 903-913.
[65] Khoury MJ, Coates RJ and Evans JP. Evidence-based classification of recommendations on use of genomic tests in clinical practice: dealing with insufficient evidence. Genet Med 2010; 12: 680-683.
[66] Bullock JG, Green DP and Ha SE. Yes, but what's the mechanism? (don't expect an easy answer). J Pers Soc Psychol 2010; 98: 550.
[67] James LR. The unmeasured variables problem in path analysis. J Appl Psychol 1980; 65: 415.

## Appendix 1: LD indices

| Unidimensional Model LD Statistics | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Item | Chi2 | x1 | x2 | x3 | x4 | x5 | x6 | x7 | x8 | x9 | x10 | x11 | x12 | x13 | x14 | x15 | x16 | x17 | x18 | x19 | x20 | x21 | x22 |
| x1 | 2.4 | | | | | | | | | | | | | | | | | | | | | | |
| x2 | 1.7 | 4.0n | | | | | | | | | | | | | | | | | | | | | |
| x3 | 1.1 | 3.5n | 4.8p | | | | | | | | | | | | | | | | | | | | |
| x4 | 1.4 | 2.1p | 1.0p | 4.7p | | | | | | | | | | | | | | | | | | | |
| x5 | 0.8 | 1.6n | 3.2p | 5.2p | 1.1p | | | | | | | | | | | | | | | | | | |
| x6 | 0.8 | 4.4n | 5.7p | 3.9p | 1.3p | 7.9p | | | | | | | | | | | | | | | | | |
| x7 | 1.5 | 2.3n | 3.7p | 7.1p | 1.3p | 9.7p | 5.8p | | | | | | | | | | | | | | | | |
| x8 | 1.6 | 6.2p | 4.9n | 4.8n | 1.6n | -0.1n | 2.8n | 2.6n | | | | | | | | | | | | | | | |
| x9 | 1.0 | 4.8p | 2.0n | 5.1n | 1.2n | 2.1n | 4.3n | 5.0n | 6.2p | | | | | | | | | | | | | | |
| x10 | 1.6 | 7.2p | 6.9n | 2.5n | 1.4n | 3.5n | 0.3n | 4.7n | 1.9p | 5.4n | | | | | | | | | | | | | |
| x11 | 3.1 | 4.5n | 6.4p | 2.4n | 1.9n | 1.2n | 2.7p | 0.5n | -0.3n | 5.0n | 4.8n | | | | | | | | | | | | |
| x12 | 1.7 | 6.1n | 4.7p | 1.5p | 1.6n | 0.7n | 2.6n | 6.7n | 7.0n | 5.9n | 3.0n | 8.2p | | | | | | | | | | | |
| x13 | 2.0 | 4.1n | 3.3p | 2.8p | 2.3n | 1.2n | 5.9p | 6.0n | 6.2n | 3.7n | 3.4n | 4.4n | 5.3p | | | | | | | | | | |
| x14 | 3.5 | 2.9p | 4.9n | 5.8n | 2.6n | 7.3n | 4.1n | 3.1n | 5.0p | 6.6p | 3.9p | 5.8n | 5.2n | 5.2n | | | | | | | | | |
| x15 | 1.1 | 4.1n | 2.2n | 4.1p | 2.1p | 9.5p | 7.4p | 13.6p | 5.1n | 2.4n | 3.8n | 1.5n | 2.5n | 4.1p | 8.2n | | | | | | | | |
| x16 | 1.6 | 2.8n | 2.8n | 4.3p | 0.8n | 8.5p | 3.0p | 6.0p | 3.1n | 5.3n | 4.6n | 2.9n | 3.0n | 4.7p | 5.3n | 15.7p | | | | | | | |
| x17 | 3.4 | 6.7n | 5.0n | 2.6n | 5.2n | 1.4n | 2.4n | 3.8n | 8.1p | 4.8n | 2.9n | 5.4p | 3.4p | 3.6n | 4.1n | 2.8n | 4.1n | | | | | | |
| x18 | 2.9 | 4.5p | 8.4n | 7.9n | 2.5n | 1.9n | 7.4n | 4.2n | 3.6n | 4.5p | 6.3p | 6.8n | 5.9n | 6.4n | 7.4p | 6.7n | 6.2n | 4.1n | | | | | |
| x19 | 1.9 | 3.8p | 2.0n | 8.7n | 1.5n | 4.5n | 1.2n | 7.5n | 9.3p | 10.6p | 8.0p | 1.3n | 4.1n | 5.1n | 13.6p | 5.1n | 6.0n | 4.6n | 5.8p | | | | |
| x20 | 3.2 | 4.2n | 5.2n | 4.3n | 2.6n | 3.2n | 1.5n | 3.1n | 3.4n | 3.8p | 3.9n | 3.0p | 3.4n | 5.2n | 4.3n | 2.0n | 5.3n | 4.9p | 8.3n | 4.0n | | | |
| x21 | 1.8 | 5.3n | 2.3n | 3.5n | 2.7n | 0.3n | 4.3n | 5.0n | 2.5n | 3.1n | 3.5n | 4.4p | 6.1p | 2.7n | 9.2n | 0.3n | 1.8n | 7.4p | 9.3n | 7.4n | 4.6p | | |
| x22 | 3.2 | 3.8n | 3.6n | 2.3n | 0.5n | 1.6n | 5.0n | 5.2n | 5.6p | 2.9n | 2.9n | 6.3n | 3.2n | 1.6n | 6.7n | 1.3n | 7.4n | 9.2p | 5.8n | 9.1n | 7.1n | 5.9p | |
| Multidimensional correlated Model LD Statistics | | | | | | | | | | | | | | | | | | | | | | | |
| Item | Chi2 | x1 | x2 | x3 | x4 | x5 | x6 | x7 | x8 | x9 | x10 | x11 | x12 | x13 | x14 | x15 | x16 | x17 | x18 | x19 | x20 | x21 | x22 |
| x1 | 2.5 | | | | | | | | | | | | | | | | | | | | | | |
| x2 | 1.8 | 4.0n | | | | | | | | | | | | | | | | | | | | | |
| x3 | 0.9 | 3.4n | 5.7p | | | | | | | | | | | | | | | | | | | | |
| x4 | 1.3 | 2.0p | 0.7p | 5.5p | | | | | | | | | | | | | | | | | | | |
| x5 | 0.7 | 1.7n | 2.9p | 0.8p | 0.8p | | | | | | | | | | | | | | | | | | |
| x6 | 1.0 | 4.7n | 5.9p | 0.9p | 1.5p | 3.1p | | | | | | | | | | | | | | | | | |
| x7 | 1.3 | 2.4n | 4.0p | 7.2n | 1.6p | 3.4p | 3.6n | | | | | | | | | | | | | | | | |
| x8 | 1.7 | 5.2p | 6.7n | 4.4n | 1.5n | 0.8n | 3.2n | 2.5n | | | | | | | | | | | | | | | |
| x9 | 1.0 | 5.1p | 2.3n | 3.9n | 1.3n | 2.7n | 3.8n | 4.5n | 6.3p | | | | | | | | | | | | | | |
| x10 | 1.6 | 7.4p | 7.3n | 1.9n | 1.4n | 4.0n | 0.3n | 4.6n | 1.9p | 5.4n | | | | | | | | | | | | | |
| x11 | 3.1 | 5.2n | 5.5p | 2.6n | 2.2n | 1.7n | 2.9p | 0.7n | 0.0n | 5.7n | 4.8n | | | | | | | | | | | | |
| x12 | 1.7 | 6.8n | 4.1p | 2.0p | 1.5n | 1.5n | 2.8n | 7.3n | 8.5n | 6.2n | 3.0n | 7.5p | | | | | | | | | | | |
| x13 | 1.5 | 2.6n | 4.6p | 1.9n | 3.0p | 3.6n | 3.7n | 8.8n | 4.4n | 2.5n | 3.3p | 5.4p | 7.9p | | | | | | | | | | |

| | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| x14 | 3.4 | 3.1p | 5.2n | 5.0n | 2.4n | 7.9n | 3.4n | 2.5n | 5.2p | 6.4n | 2.7p | 5.9n | 5.3n | 6.0p | | | | | | | | |
| x15 | 1.2 | 5.6n | 2.1n | 1.8n | 2.1n | 1.7p | 4.1n | 11.8n | 7.2n | 2.8n | 4.1n | 1.9n | 3.0n | 2.9n | 9.1n | | | | | | | |
| x16 | 1.3 | 3.8n | 3.4n | 5.0n | 0.8n | 1.1p | 1.3n | 3.7n | 3.9n | 5.0n | 4.4n | 3.4n | 3.2n | 4.6n | 5.1n | 5.6p | | | | | | |
| x17 | 3.4 | 8.1n | 5.9n | 2.4n | 5.7n | 2.9n | 2.5n | 4.1n | 7.8p | 5.3n | 3.2n | 4.7p | 3.1n | 2.7n | 4.4n | 3.9n | 4.6n | | | | | |
| x18 | 2.2 | 8.3p | 4.3n | 8.2n | 1.0p | 3.2n | 9.5n | 4.3n | 5.6p | 11.4p | 12.0p | 4.4p | 3.9p | 5.7n | 16.5p | 8.6n | 6.7n | 5.1p | | | | |
| x19 | 2.3 | 4.3n | 3.2n | 7.0n | 1.9n | 6.1n | 0.8n | 7.5n | 10.0p | 10.0n | 6.6p | 1.8n | 4.5n | 4.4p | 9.6p | 6.6n | 6.7n | 5.3n | 14.5p | | | |
| x20 | 2.8 | 3.7n | 4.8n | 3.5n | 2.6n | 2.8n | 1.5n | 2.7n | 3.5p | 4.3n | 4.3n | 3.9p | 3.9p | 8.3p | 5.4n | 2.1p | 5.7p | 5.5p | 9.3p | 5.4n | | |
| x21 | 1.5 | 4.1n | 2.3p | 4.4p | 3.0p | 0.6p | 4.9p | 5.1p | 2.9n | 3.5n | 3.9n | 5.4p | 7.5p | 7.3p | 10.0n | 0.5p | 3.0p | 9.1p | 7.4p | 9.8n | 5.1p | |
| x22 | 3.2 | 4.2n | 3.8n | 2.5n | 0.7n | 1.5n | 5.8n | 5.3n | 5.7n | 2.9n | 2.9n | 7.5n | 3.5n | 1.5n | 6.6n | 1.7n | 8.4n | 9.2p | 7.0p | 10.2n | 6.9n | 6.7p |

Bifactor Model LD Statistics

| Item | Chi2 | x1 | x2 | x3 | x4 | x5 | x6 | x7 | x8 | x9 | x10 | x11 | x12 | x13 | x14 | x15 | x16 | x17 | x18 | x19 | x20 | x21 | x22 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| x1 | 2.2 | | | | | | | | | | | | | | | | | | | | | | |
| x2 | 1.6 | 5.0p | | | | | | | | | | | | | | | | | | | | | |
| x3 | 0.9 | 3.4n | 6.2p | | | | | | | | | | | | | | | | | | | | |
| x4 | 1.3 | 2.0p | 1.1p | 5.6p | | | | | | | | | | | | | | | | | | | |
| x5 | 0.4 | 1.6p | 4.3p | 0.5p | 2.0p | | | | | | | | | | | | | | | | | | |
| x6 | 1.0 | 3.9n | 7.3p | 1.0p | 1.9p | 3.0p | | | | | | | | | | | | | | | | | |
| x7 | 1.2 | 2.6p | 4.6p | 7.2n | 1.8p | 2.6p | 3.7n | | | | | | | | | | | | | | | | |
| x8 | 1.8 | 2.7p | 3.8n | 5.4n | 1.7n | 0.3n | 2.6n | 2.3n | | | | | | | | | | | | | | | |
| x9 | 1.0 | 4.8p | 2.3n | 4.5n | 1.3n | 1.4n | 3.4n | 4.5n | 5.2p | | | | | | | | | | | | | | |
| x10 | 1.4 | 7.6p | 7.1n | 1.9n | 1.3n | 2.5n | 0.4n | 4.5n | 1.9p | 5.5n | | | | | | | | | | | | | |
| x11 | 3.2 | 4.1n | 3.5p | 2.9n | 1.9n | 1.9p | 3.8p | 1.1p | -0.0p | 6.4n | 4.9n | | | | | | | | | | | | |
| x12 | 1.9 | 4.7n | 3.2n | 2.1p | 1.7n | 0.3n | 3.5p | 2.6n | 5.3n | 7.5n | 3.1n | 6.5n | | | | | | | | | | | |
| x13 | 1.9 | 3.9n | 3.8p | 2.1p | 2.4p | 2.6n | 4.3p | 7.7n | 7.2n | 3.1n | 3.2n | 4.3p | 6.2p | | | | | | | | | | |
| x14 | 3.0 | 3.3p | 4.7n | 5.1n | 2.2n | 6.2n | 2.9n | 2.3n | 5.6p | 6.5n | 2.2n | 6.0n | 5.7n | 5.3p | | | | | | | | | |
| x15 | 0.7 | 3.2n | 2.8p | 2.0n | 2.9p | 1.6n | 4.6n | 6.0n | 4.3n | 1.6n | 2.6n | 2.4p | 3.5p | 3.5p | 6.6n | | | | | | | | |
| x16 | 1.0 | 2.5n | 2.7n | 5.2n | 1.0p | 0.5n | 1.8n | 4.2n | 3.1n | 4.7n | 4.2n | 3.5n | 3.5n | 5.0n | 4.8n | 4.1p | | | | | | | |
| x17 | 3.7 | 5.9n | 5.9n | 2.6n | 5.5n | 1.4n | 2.0n | 3.9n | 8.2p | 5.6n | 3.3n | 5.0p | 3.6n | 3.9n | 4.5n | 2.6n | 4.3n | | | | | | |
| x18 | 2.9 | 4.7p | 7.6n | 7.3n | 1.8n | 2.2n | 7.0n | 4.0n | 4.5n | 5.2p | 7.5p | 6.4n | 6.2n | 6.0n | 9.2p | 7.5p | 7.3p | 4.2n | | | | | |
| x19 | 1.7 | 4.6n | 1.9n | 7.0n | 1.5n | 2.9n | -0.2n | 6.6n | 11.5p | 10.4n | 7.4n | 1.6n | 4.7n | 4.9n | 8.7p | 3.0n | 5.3n | 5.5n | 7.6p | | | | |
| x20 | 3.1 | 5.3n | 6.2n | 4.1n | 2.7n | 1.4n | 1.6n | 2.8n | 4.0n | 3.8p | 3.8n | 2.8p | 3.8n | 5.4p | 4.1n | 2.2p | 5.5n | 4.3n | 8.1n | 4.4n | | | |
| x21 | 1.8 | 7.0n | 2.8n | 3.6n | 2.7n | -0.0n | 4.8n | 5.4n | 4.4n | 3.1n | 2.6n | 4.6n | 6.9n | 2.9n | 8.0n | 0.6p | 2.2p | 6.6p | 9.0n | 7.0n | 3.4p | | |
| x22 | 3.3 | 4.7n | 3.7n | 2.5n | 0.5n | 2.0p | 5.5n | 4.8n | 7.3n | 3.1n | 2.8n | 6.8n | 4.2n | 1.5n | 6.4n | 1.3n | 7.5n | 9.3p | 5.6n | 9.8n | 5.9n | 3.9p | |

## Appendix 2: Item level indices

Item Fit Statistics

| | Bifactor Item fit statistics | | | | | Multidimensional 2nd order | | | | | Multidimensional Correlated | | | | | Unidimensional | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| item | p | k | t (adj alpha) | sig/ns | S_X2 | S_X2 | p | k | t (adj alpha) | sig/ns | S_X2 | p | k | t (adj alpha) | sig/ns | S_X2 | p | k | t (adj alpha) | sig/ns |
| x1 | 1.44E-12 | 35 | 0.080 | ns | 128.3 | 226.7 | 6.4E-30 | 35 | 0.080 | ns | 135.9 | 4E-14 | 34 | 0.077 | ns | 266.0 | 3.2E-43 | 23 | 0.052 | sig |

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| x2 | 1.51E-09 | 35 | 0.080 | ns | 109.1 | 96.2 | 9.9E-07 | 39 | 0.089 | ns | 97.7 | 3E-08 | 33 | 0.075 | ns | 144.5 | 3.1E-16 | 32 | 0.073 | ns |
| x3 | 5.57E-15 | 35 | 0.080 | ns | 142.9 | 129.8 | 6.7E-11 | 42 | 0.095 | ns | 137.7 | 2E-15 | 31 | 0.070 | ns | 168.5 | 5.2E-22 | 28 | 0.064 | ns |
| x4 | 2.12E-09 | 39 | 0.089 | ns | 114.8 | 112.2 | 4.7E-07 | 48 | 0.109 | ns | 115.9 | 8E-10 | 38 | 0.086 | ns | 152.3 | 5.9E-18 | 31 | 0.070 | ns |
| x5 | 3.71E-16 | 33 | 0.075 | ns | 146.0 | 257.1 | 3.5E-35 | 36 | 0.082 | ns | 157.4 | 4E-18 | 33 | 0.075 | ns | 333.6 | 3.8E-55 | 26 | 0.059 | ns |
| x6 | 1.35E-12 | 35 | 0.080 | ns | 128.4 | 137.6 | 2.3E-12 | 41 | 0.093 | ns | 130.3 | 2E-14 | 30 | 0.068 | ns | 178.3 | 4.2E-25 | 25 | 0.057 | ns |
| x7 | 1.09E-27 | 31 | 0.070 | ns | 205.8 | 308.0 | 2.6E-41 | 44 | 0.100 | ns | 220.8 | 1E-27 | 38 | 0.086 | ns | 183.3 | 9.6E-23 | 33 | 0.075 | ns |
| x8 | 7.32E-49 | 23 | 0.052 | ns | 294.0 | 473.0 | 9.8E-82 | 29 | 0.066 | ns | 419.9 | 3E-73 | 25 | 0.057 | ns | 243.7 | 1.5E-33 | 34 | 0.077 | ns |
| x9 | 5.54E-29 | 24 | 0.055 | ns | 196.3 | 159.2 | 3.8E-16 | 40 | 0.091 | ns | 161.5 | 1E-20 | 28 | 0.064 | ns | 154.9 | 7.8E-15 | 42 | 0.095 | ns |
| x10 | 8.14E-24 | 31 | 0.070 | ns | 184.9 | 172.1 | 1.1E-17 | 42 | 0.095 | ns | 190.3 | 1E-23 | 34 | 0.077 | ns | 172.8 | 4.1E-18 | 41 | 0.093 | ns |
| x11 | 2.77E-25 | 35 | 0.080 | ns | 201.5 | 163.8 | 5.5E-16 | 43 | 0.098 | ns | 186.4 | 6E-23 | 34 | 0.077 | ns | 150.2 | 2.9E-15 | 38 | 0.086 | ns |
| x12 | 2.91E-23 | 31 | 0.070 | ns | 181.8 | 161.5 | 1.7E-17 | 37 | 0.084 | ns | 176.8 | 1E-22 | 30 | 0.068 | ns | 136.2 | 5.5E-13 | 38 | 0.086 | ns |
| x13 | 4.76E-26 | 31 | 0.070 | ns | 197.0 | 145.2 | 5.9E-12 | 47 | 0.107 | ns | 145.4 | 2E-15 | 35 | 0.080 | ns | 166.5 | 1.9E-16 | 43 | 0.098 | ns |
| x14 | 5.64E-10 | 36 | 0.082 | ns | 113.6 | 132.3 | 5.1E-11 | 43 | 0.098 | ns | 130.0 | 4E-13 | 34 | 0.077 | ns | 117.2 | 3E-10 | 37 | 0.084 | ns |
| x15 | 7.8E-11 | 32 | 0.073 | ns | 112.2 | 114.2 | 1.3E-08 | 42 | 0.095 | ns | 123.7 | 2E-12 | 33 | 0.075 | ns | 126.5 | 1.2E-10 | 41 | 0.093 | ns |
| x16 | 2.11E-30 | 39 | 0.089 | ns | 238.1 | 246.9 | 1.2E-29 | 45 | 0.102 | ns | 250.4 | 8E-35 | 34 | 0.077 | ns | 246.3 | 6.5E-32 | 39 | 0.089 | ns |
| x17 | 7.5E-11 | 35 | 0.080 | ns | 117.5 | 197.1 | 7.2E-25 | 34 | 0.077 | ns | 120.6 | 3E-12 | 32 | 0.073 | ns | 127.1 | 5.3E-10 | 44 | 0.100 | ns |
| x18 | 7.16E-09 | 34 | 0.077 | ns | 103.0 | 190.9 | 2.3E-23 | 35 | 0.080 | ns | 117.3 | 8E-11 | 35 | 0.080 | ns | 114.3 | 4.4E-09 | 40 | 0.091 | ns |
| x19 | 6.1E-117 | 46 | 0.105 | ns | 695.9 | 637.1 | 5E-111 | 36 | 0.082 | ns | 611.3 | 3E-108 | 32 | 0.073 | ns | 835.1 | 1E-155 | 31 | 0.070 | ns |
| x20 | 6.81E-39 | 47 | 0.107 | ns | 301.8 | 242.3 | 5.2E-26 | 53 | 0.120 | ns | 259.4 | 6E-32 | 45 | 0.102 | ns | 249.4 | 1.1E-31 | 41 | 0.093 | ns |
| x21 | 1.03E-87 | 44 | 0.100 | ns | 545.5 | 1057.5 | 1E-198 | 36 | 0.082 | ns | 741.4 | 4E-128 | 43 | 0.098 | ns | 1217.7 | 7E-234 | 34 | 0.077 | ns |
| x22 | 1.95E-95 | 41 | 0.093 | ns | 575.5 | 947.7 | 1E-175 | 36 | 0.082 | ns | 848.8 | 2E-149 | 44 | 0.100 | ns | 1439.0 | 4E-279 | 36 | 0.082 | ns |