ISSN: 2584-0355
DOI: doi.org/10.51219/MCCRJ/Satyendra-Nath-Chakrabartty/02

URF PUBLISHERS
connect with research world

Medical & Clinical Case Reports Journal

https://urfpublishers.com/journal/case-reports

Vol: 1 & Iss: 1

Research Article

Drug Addiction and Quality of Life: Methodological Issues

Satyendra Nath Chakrabartty

Indian Statistical Institute, Indian Maritime University, Indian Ports Association, India

***Corresponding author:** Chakrabartty S. N, Indian Maritime University, Indian Ports Association, India, Tel: 919831597909; Email: chakrabarttysatyendra3139@gmail.com. DOI: doi.org/10.51219/MCCRJ/Satyendra-Nath-Chakrabartty/02

Citation: Chakrabartty S. N. (2023). Drug Addiction and Quality of Life: Methodological Issues. Medi Clin Case Rep J 1(1), 5-11.

Received: 25 April, 2023; **Accepted:** 06 June, 2023; **Published:** 09 June, 2023

A B S T R A C T

Objective: Substance abuse research data containing categorical and continuous variables often violate assumptions of parametric statistical methods. Clustering of individuals, lack of repeated measurements, missing data, non-representative samples, etc. aggravates the problem. Quality of life (QoL) measures suffer from meaningful application of statistical methods. The paper describes statistical approaches which fit well with structure of drug addiction data and measures of QoL, facilitating better analysis and interpretations of results.

Method: Converting ordinal item scores to normally distributed continuous scores in the range [1-100], irrespective of number of response-category in items. Such transformations fit well with structure of drug addiction data and measures of QoL and facilitate better analysis and interpretations of results.

Results: Proposed method enables parametric statistical analysis leading to meaningful comparisons and inferences, finding equivalent scores, computation of responsiveness of the scale i.e., ability to assess changes across time and psychometric qualities like reliability, as per definition, Factorial validity reflecting the main factor for which the test was developed.

Conclusions: Considering theoretical advantages, the proposed method generating normally distributed scores is recommended. Future studies with longitudinal data suggested finding sensitivity with emphasis on progression of disease and to different therapeutic interventions.

Keywords: Contingency Table; Factorial Validity; Logistic regression; Normal distribution; Parallel tests; Quality of life; Substance abuse; Theoretical Reliability

Introduction

Drugs of different types are being used increasingly to relieve pain and enjoy "Feel-good" effects, alleviate stress, cope with mental and emotional pain, alter or avoid reality. Some drugs act brain and alter mood in ways which are not approved socially like feeling of euphoria, restlessness, nightmare, hallucinations and delusions [1].

Drug addiction (substance use disorder) is a global disease associated with significant morbidity, mortality and is characterized by compulsive and uncontrollable drug cravings. Drugs bring changes in behaviors by affecting brain and central nervous system which in turn produce physical, psychological, behavioral changes and harm the social fabric of people's lives and social relationships [2]. However, different types of drugs affect differently on one's brain and Quality of Life (QoL) of substance users. Need for specific questionnaires covering substance use, misuse and dependency and their effects on QoL was felt [3].

Using logistic regression [4] identified five factors for addiction viz. religion, employment status, previous family history, accessibility of drug and peer pressure. Unsanctioned psychoactive drugs including cannabis accounts for an estimated 83 disability-adjusted life years lost per 100 000 population [5].

Attempts made to estimate prevalence and addiction-related harms, number of drug addicts at global and national levels. However, due to poor quality of data, comparisons of countries

and regions across time and space are rather indicative. Major sources of primary data are websites of the World Health Organization (WHO), the United Nations Office on Drugs and Crime (UNODC), Institute for Health Metrics and Evaluation, Alberta Gambling Research Institute, etc. Need for standardized and rigorous methods for data collection, and reporting have been highlighted for better assessment in substance use and addiction burden [6]. Substance abuse research data containing categorical and continuous variables, clustering of individuals in selected areas, significant missing data and outliers, non-representative samples, lack repeated measurements. Even the independent observations assumption may be violated when data are collected from participants who are clustered within larger units, say schools, clinics, communities. Such cluster-wise dependency may not result in independent observations [7].

Thus, statistical analysis of addiction data and the relevant factors need to be innovative for answering questions how, why, and for whom prevention programs are effective. For example, error scores may not follow normal distribution when the dependent variable is binary. Here, logistic regression could be useful where the regression coefficients are the logarithm of the odds ratio.

The paper aims at describing statistical approaches which fit well with structure of drug addiction data and measures of QoL and facilitate better analysis and interpretations of results.

Analysis of Drug Addiction Data

Frequency based approaches

Statistical approaches based on frequencies includes selection of dependent variable (Y) drug addiction (like Peer pressure (yes – no), accessibility of drug and a set of independent variables which may cause drug abuses and conducting Chi-square test of independence by $X^2 = \sum_{i=1}^{n} = \frac{(observedvalue_i - Expectedvalue_i)^2}{Expectedvalue_i}$ where expected value of *(i-j)* cell= $\frac{i-thRowtotal * j-thcolumntotal}{GrandTotal}$

To investigate gender effect on drug use, consider the following 2 × *2* contingency table with sample size *N*:

		Drug addiction		*Total*
		Positive (Success)	Negative (Failure)	
Gender	Male	f_{11}	f_{12}	Row total (R_1)
	Female	f_{21}	f_{22}	Row total (R_2)
Total		Column total (C_1)	Column total (C_2)	Grand Total (N)

Table 1: Contingency Table – Gender vs. Drug addiction.

Measures of association considering cell frequencies of contingency table are:

- Chi-square measure of association $X^2 = \sum \frac{(f_{observed} - f_{Expected})^2}{f_{observed}} \sim X^2$ distribution with (*r*-1) (*s*-1) degrees of freedom (df) for *r*×*s* contingency table
- Pearson's Contingency Coefficient $c = \sqrt{\frac{x^2}{N + X^2}}$ to measure relative (strength) of an association between two variables where $0 \le v \le 1$. The measure can be used for incidence and prevalence studies
- Cramer's V-Coefficient (V): X^2 may get increased by large sample size, even if the variables may not have any substantive relationship. Cramer's V-Coefficient improves association by $v = \frac{\sqrt{X^2}}{N^{(q-1)}}$ were

  q = minimum (r, c) and $0 \le v \le 1$

  Hypothesis frequently used in measures of association is

  H_0: The two attributes are independent (no significant association between the two attributes) is tested against the alternate hypothesis

  H_1: The two attributes are dependent (there is significant association between the two attributes)

  X^2 measure of association, C-coefficient (C), V-coefficient, etc. can be applied to find association with clinical findings related to the drug addiction along with testing of significance of association.

  Here, proportion of success are $z = \frac{p_1 - P_2}{\sqrt{p(1-p)(\frac{1}{R_1} + \frac{1}{R_2})}}$

  Let population proportions are $\prod_1 - \prod_2$

  The z-statistic for testing $H_0 : \prod_1 = \prod_2$ for large/moderate sample size is

$$z = \frac{p_1 - P_2}{\sqrt{p(1-p)(\frac{1}{R_1} + \frac{1}{R_2})}} \sim N(0,1) \text{ where } p = \frac{f_{11} + f_{21}}{N}$$

For small sample, Fisher's exact test is preferred to test equality of population proportions

Wald Confidence interval (CI) for $(\prod_1 - \prod_2)$ at 95% level is

$$w = (\overline{p_1} - \overline{p_2}) \pm 1.96 \frac{\sqrt{\overline{P_1}(1-\overline{P_1})}}{R_1 + 2} + \frac{\overline{P_2(1-\overline{P_2})}}{R_2 + 2}$$

For small samples, Wald CI for $(\prod_1 - \prod_2)$ is improved by considering $\overline{P_1} = \frac{f_{11} + 1}{R_1 + 2}$ $\overline{P_2} = \frac{f_{21} + 1}{R_2 + 2}$ $\sqrt{\frac{P_1(1-P_1)}{R_1} + \frac{p_2(1-p_2)}{R_2}}$ in each of the two samples and taking Agresti-Caffo-CI for $(\prod_1 - \prod_2)$ at 95% level as $(\overline{p_1} - \overline{p_2})$ $\pm 1.96 \frac{\sqrt{\overline{P_1}(1-\overline{P_1})}}{R_1 + 2} + \frac{\overline{P_2(1-\overline{P_2})}}{R_2 + 2}$

The Agresti-Caffo CI is closer to the nominal level than the Wald-CI. Thus, Agrestic-Caffo-CI is preferred. 2x2 table can also help to find prevalence of exposure by

Odd Ratio (*OR*) = $\frac{f_{11} * f_{22}}{f_{12} * f_{21}}$ *OR Relative Risk* $\frac{\prod_1}{\prod_2}$ which is estimated by $\frac{p_1}{p_2}$ *OR* compares categorical outcome data where OR = RR $(\frac{1-\prod_2}{1-\prod_1})$ of exposure between two groups formed by outcome and *RR* compares incidence of disease between two groups formed by outcome. Both fail if the assumption of independence is violated. CI of *OR* and *RR* can be formed to reflect range of uncertainty. *OR* compares categorical outcome

data where $(\frac{1-\prod_2}{1-\prod_1})$ Sampling distribution for sample RR $(\frac{p_1}{p_2})$ is highly skewed.

2 × 2 Contingency table can be extended to *r-rows* × *s-columns* to accommodate levels of drug addictions say mild, moderate, severe, etc. and various level of income or age-group. However, such associations depend heavily on the sample characteristics. Predominance of one gender or number of drug addicts may distort the result.

Correlations

Point bi-serial correlation is suitable to find correlation between categorical variable (Y) (say drug addict or not) and a continuous variable (X) (say age, monthly income, etc.) and is given by $r_{pb} = \frac{M_1 - M_0}{S_n}\sqrt{\frac{n_1 n_2}{n}}$

$M_1 : \overline{X} or\ the\ group\ with\ Y = 1\ (sample\ size n_1)$

$M_0 : \overline{X} for\ the\ group\ with\ Y = 0\ (sample\ size n_2)\ \ n = n_1 + n_2$

$S_n = \sqrt{\frac{1}{n}\Sigma(X_i - \bar{X})^2}$ and $-1 \le r_{pbs} \le 1$ For H_0: $r_{pbs} = 0$, the test statistics is $t = \sqrt{\frac{n_1 - n_2 - 2}{1 - r_{pb}^2}} \sim t$ distribution with $(n_1 - n_2 - 2)$ df.

Logistic Regression (LR)

LR is used when the dependent variable is qualitative or categorical. For categorical dependent variable, several assumptions of the ordinary least square are violated. Qualitative response variable is either binary or have multiple categories. The linear logistic regression model is given by:

$ln\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots \ldots \ldots \ldots + \beta_k X_k$ where $0 \le \pi \le 1$ is the probability of success and $(1 - \prod)$: the probability of failure

β_0: Constant term

β_i: Coefficients of *i*-th independent variables, *i*=1, 2, 3, k

$X_i : i - th$ independent variable $\forall i = 1, 2, 3..k$

Here, $\frac{prohabilityofsucess}{prohabilityoffailure} = \frac{\prod}{1-\prod}$ is odd ratio of success and (ii) Exp (βj) where *j* = 1, 2… k is a factor by which the odds of occurrence of success change by a unit increase in the *j*-th independent variable.

Test of individual LR regression coefficient i.e., $H_0 : \beta_m = 0$ against $H_1 : \beta_m \neq 0$ is done by the Wald test by $\frac{\widehat{\beta_m}}{\text{standard error of } \widehat{\beta_m}}$ where standard error (SE) of $\widehat{\beta_m} = \sqrt{\hat{V}(\widehat{\beta_m})}$ $and z \sim$ N(0,1) asymptotically. Equivalently, $W = \left(\frac{\widehat{\beta_m}}{\sqrt{\hat{V}(\widehat{\beta_m})}}\right)^2$ can be used where $w \sim x^2$ distribution approximately H_0. Rejection of $H_0 \Rightarrow X_i$ adds something to the model.

One can use CI for the odds ratio to determine whether odds ratio = 1. If the CI does not contain one, then we conclude that the odds ratio is statistically significant. For small samples, reliability of Wald statistics is questionable [4]. However, LR method is sensitive to dependent variables and the researcher need to choose them correctly.

Multilevel analysis

Intra-class correlation is used to evaluate cluster-based dependency data [8]. However, both positive and negative intra-class correlation may influence Type I error (unjust rejection of). Multilevel analysis enabling simultaneous analysis of several levels of data can better analyze clustered data for cluster data, multilevel model includes individuals at Level 1 and schools (say) at Level 2. Relationships between *i*-th individual and *j*-th school are given by the following three equations:

Individual (Level 1): $Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + \epsilon_{ij}$ (1)

School (Level 2) $\beta_{oj} = \gamma_{00} + \gamma_{01} W_j + u_{oj}$ (2)

$\beta_{1j} = \gamma_{10} + u_{1j}$ (3)

The three equations show relationship between individual-level predictors *X*-matrix (gender, age, other relevant covariates) and school-level predictors in the *W* matrix (like assignment to program or control groups or other school characteristics). Estimation of error terms $\in_{ij}, u_{oj}, u_{ij}$ in (1,2,3) respectively allow non-zero intra-class correlations which are incorporated in the analysis. In addition, multilevel models can also be used to study effects at the cross-levels (say, school effect on individuals) or third order cross-levels *School Individual Classroom*. However, applications of higher order cross-levels are rare for drug prevention studies [9]. One limitation of multilevel analysis is weakened causal inference without randomization. However, there could be situations where randomization is either not possible or may be unethical [10-11] studied *Fighting Back* program focusing on the use of multiple comparison sites for each treatment community where only respondents were selected randomly and not the cities or communities or sites. In addition, coding for each variable is important, since intercepts and slopes at each level are explained at the next highest level. If intercepts or slopes have no useful interpretation, then it is not meaningful to explain them at higher levels. Multilevel analysis has been used to substance abuse prevention and cross-site evaluation of CSAP's high-risk youth programs CSAP's community partnerships [7].

Covariance Structure Analysis (CSA)

To know the extent to which predictors differ across subgroups (gender, income, racial groups, etc.), CSA are used to test the equality or invariance of effects across groups, using test comparing the model with and without parameters freed across groups. Disadvantages of CSA are needed to check assumption of multivariate normal distribution and computation of fourth order moments.

Asymmetric data

Drug addiction data are often skewed and non-normal. Computer-intensive methods like bootstrap, randomization tests, etc. use observed data and finds the significance of an effect without making assumptions about underlying distributions and thus, help in analysis of substance abuse research data. However, such computer-intensive approaches have not been widely used in substance abuse data.

Quality of Life

People with substance use generally have lower QoL-scores than those without such problems [12]. Substance use does not have enduring benefits for subjective feelings of wellbeing [13]. Like drug addiction data, QoL measures suffer from problems relating to meaningful application of statistical analysis. In the area of substance abuse, QoL has been evaluated covering among others functioning, well-being and life satisfaction [14-; 15] Studies to find relationship between QoL, treatment and substance use have mostly considered habits of consuming alcohol with few studies reporting on drug dependent samples. While abstinence alone may have little effect on QoL [16]. gain on QoL may accrue gradually with increasing length of abstinence exceeding the initial six months [17] Investigation of QoL in the addiction field is in its infancy and further works are suggested [18].

QoL for substance users have been evaluated using generic instruments like 36-Item Short Form Health Survey questionnaire (SF-36), World Health Organization Quality of Life Assessment-BREF (WHOQOL-BREF), specific tool like Drug Users Quality of Life Scale (DUQOL) [19-20].

Psychometric properties of SF-36 and WHOQOL-BREF emerging from various populations may not be applicable to substance users. Moreover, scales for assessing QoL differ with respect to number of items, item formats, scoring methods and are not comparable [21]. felt need of clarifications of WHOQOL-BREF.

In SF 36, original item scores are rescaled to range between 0 to100. A high score indicates higher favorable health-state. However, total SF-36 scores are not defined due to several dimensions being measured by the scale **(**http://www.webcitation.org/6cfeefPkf**).** Descriptive statistics, reliability, validity, etc., are obtained separately for each sub-scale. Mean, SD and shape of distribution, reliability, validity, and discriminating power are different for Yes-No type, *K*- point scales for *K*= 3, 5, 6-point [22]. Studies to investigate factor structure of SF-36 using factor analysis [23]. structural equation model [24]. confirmed multidimensional structure of the SF-36 [25]. found large number of articles reporting calculation of total SF-36 scores by different ways including arithmetic average of eight sub-classes.

(Table 2) describes major features of QoL scales used for substance users

Scale	Sub-scales (Dimensions)	No. of levels (*K*)	Remarks
SF 36 (36 items)	Physical functioning	*K* =3	1 is recorded as 0 and 3 is recorded as 100
	Role limitations due to physical health	*K* =2: Binary	1 is recorded as 0 and 2 is recorded as 100
	Role limitations due to emotional problems	*K* =2: Binary	1 is recorded as 0 and 2 is recorded as 100
	Energy/ Fatigue	*K* =6	For item 23, 27 (1:100 and 6: 0), requiring reverse scoring to ensure higher score □ better QoL. For item 29, 31(6: 100 and 1: 0)
	Emotional well-being	*K* = 6	Item 24, 25, 26, 28, 30 (6: 100 and 1: 0) Item 26 (1:100 and 6: 0) – requires reverse scoring
	Social functioning	*K* = 5	Item 20 (1:100 and 5:0) requires reverse scoring and Item 32 (1: 0 and 5:100)
	Pain	*K* =6 for Item 21 & *K* =5 for Item 22	21 (1: 100 and 6:0) and 22 (1:100 and 5: 0)
	General health	*K* =5	Item 1, 34, 36 (1:100 and 5:0). For item 33, 35 (1:0 and 5:100)
WHOQOL-BREF (26-items)	Physical health	7 items, *K* =5	Computation of Domain scores (1) reverse scoring two items of physical domain and one item of psychological domain, (2) using the mean item-score of each domain and (3) multiplying the mean scores by 4 to get domain scores, Domain scores are transformed linearly to a 0-100 scale.
	Psychological health	6 items, *K* =5	
	Social relationships	3 items, *K* =5	
	Environmental health	8 items, *K* =5	
	General health	2 items, *K* =5	
DUQOL (22-items)		*K*=7 for each item	Item score ranges from 1 (very unsatisfied) to 7 (very satisfied). Here higher score indicates higher satisfaction with quality of life.

Table 2: Major features of illustrative QoL scales.

Observations

Generated data are ordinal and discrete.

Assume constant distance between two successive levels of an item (Equidistance property). For a 7-point item, equidistant property requires constant value of distance between *j*-th and *(j+1)*-th levels $\forall j = 1, 2, 3, 4, 5, 6$

Arithmetic averages requiring equidistant scores are not meaningful for ordinal item scores [26].

Summative scores for domain and test giving equal importance to items and domains are un-justified due to different values of inter-item correlations, item-total correlations and factor loadings of the items and domains.

Mean, SD of test tends to increase with increase in number of levels and may influence mean more than the underlying variable [27].

Distribution of scores of items, domain and test are different and skewed. If two variables X and Y follow two different distributions, $X + Y = Z$ is most meaningful if

$(p = Z = z) = p(X = X, Y = Z - X)$ for discrete case and $p(z = z) = P(X + Y \le Z) = \int_{-\infty}^{\infty}(\int_{-\infty}^{Z} f_{X,Y}(x, t-x)dt)dx$ for continuous case. Thus, it is necessary to know probability density function (pdf) of *X* and *Y* and their convolution.

PCA, FA, *t*-test, paired *t*-test, *F*-test, etc., assume normal distribution of the variables under study. Results may go wrong if assumptions of the techniques used are violated. Even, high r_{xy} may not imply linearity between *X* and *Y*. If = 1, 2, 3., 3 $r_{X,X^2} > 0.9$ and $r_{X,X^3} > 0.9$ 0.9 despite each of $X^2 X^3$, is non-linear function of *X* Linearity between *X* and *Y* can be tested by checking normality of error score by testing $H_0 : S_E^2 = 0$ where $S_E^2 = \frac{1}{n}\Sigma(Y_i - \hat{Y}_i)^2$ for a sample size *n* denotes variance of error scores

Proposed method

[28] proposed method of converting ordinal scores of *i*-th item to continuous scores satisfying equidistant property (-scores) in ratio scale and further transforming -scores linearly to proposed scores(-scores) in ratio scale and further transforming -scores linearly to proposed scores (-scores) where $1 \le P_i \le 100$ and normal distribution, irrespective of number of levels (*K*) of an item, *K*= 2, 3, 4, 5 and so on. Equidistant scores are obtained by assigning different weights to different levels of different items. Let be the weight to the *j*-th level of the *i*-th item (say 5-point) so that $w_{i1}, 2w_{i2}, 3w_{i3}, 4w_{i4,} 5w_{i5}$ forms and arithmetic progression. Clearly $j.w_{ij} - ((j-1)w_{i(j-1)}$ is constant, value of which will be different for different items. A positive value of constant will make *E*-scores monotonic. Here $w_{ij}s$ are based on empirical probabilities obtained from the basic Item score matrix. If $f_{ij} = 0$ for a particular *j*-th level of an item, the method fails and can be taken as zero value for scoring Likert items as weighted sum. Different weights to different levels of different items may break ties of individual total scores which are common in usual summative scores and thus, improve discriminating power of the test or sub-test.

E_i-scores are normalized by $z_i = \frac{E_i - \overline{E}}{SD(E)} \sim N(0,1)$ To avoid negative values, Z-scores are transformed to proposed scores $(p_i - scores) by \;\; p_i = \frac{99*(Z_{ij} - Minz_{ij})}{Maxz_{ij} - Minz_{ij}} + 1 \;\; 1 \le p_i \le 100$ follows p_i normal. Domain score is the sum of of the domain and test score is sum of all domain scores, each following normal distribution, parameters of which can be estimated from the data.

Advantages of proposed scores

Better admissibility of arithmetic average where scores are normally distributed

Measurement of total test score for each individual like SF_{Total}

Dimension score D_i and proposed scale scores are continuous, monotonic, normal and facilitate undertaking parametric analysis including estimation of population mean (μ) population variance (σ^2) confidence interval of testing statistical hypothesis like $H_0 : \mu_i = \mu_2$ or $H_0 : \sigma_1^2 = \sigma_2^2$. either for longitudinal data or snap-shot data.

P-scores reduce significantly number of tied scores. Thus, most of the individuals can be given unique ranks.

Progress path of a patient or a group of patients across time i.e., trajectories over time is analogous to hazard function and can be used to compare response to treatments from the start This is especially important for substance users who are undergoing treatments. Such trajectories can help to identify high-risk groups.

For two QoL scales X with normal pdf and Y with normal pdf one can find regression equation of the form $y = \alpha_1 + \beta_1 X$ $X = \alpha_2 + \beta_2 Y$ predict X with knowledge of Y. However, the two regression lines differ and thus, relationship between X and Y will not be unique. Better is to find equivalent score combinations (X_0, Y_0) of the two QoL scalesby solving the equation $\int_{-\infty}^{x_0} f(x)dx = \int_{-\infty}^{y_0} g(y)dy$ for a given value say i.e. area of the curve up to = area of the curve up to [29].This avoids the problems of linear equating or percentile equating. The equation can be solved using standard normal table. The method of finding equivalent score-combinations is possible even if the scales have different number of items or dimensions.

Psychometric qualities

Validity

Validity of a QoL instrument is often reported as correlation between the QoL score (X) in question and another scale (Y) which may be influenced by mismatch of dimensions covered, different score ranges, different distributions of scores of the two scales, etc. Moreover, reflects validity of X and Y also, If r_{xy} is high, need of two different scales may be questioned. Normality satisfies the assumptions of PCA enabling computation of eigenvalue and consider factorial validity as $\frac{\lambda_1}{\sum \lambda_i}$ where λ_1 is the highest eigenvalue associated with the first principal component. Factorial validity reflects *the main factor for which the test was developed* and accounts for $\frac{\lambda_i}{\sum \lambda_i} * 100$ percent of overall variability. Such factorial validity from single administration of a test avoids the problems of construct validity and is independent of criterion scale [30].

Reliability

Avoiding uni-dimensionality assumption of Cronbach alpha, [31] proposed to dichotomize a test in two parallel subtests (*g*-th and *h*-th) and to compute test reliability as per theoretical definition $r_{tt(Theoretical)} = \frac{True\ score\ variance\ (S_T^2)}{Observed\ score\ variance\ (S_X^2)}$

where Error variance S_E^2 *for Sample Size N* is

$$S_E^2 = \frac{1}{N} = [\| X_g \|^2 + \| X_h \|^2 - 2 \| X_g \| \| X_h \| \cos\theta_{gh} \quad (4)$$

where $\left\| X_g \right\|$ denotes length of the *g*-th vector $\| X_h \|$ is defined similarly and θ_{gh} is the angle between the *g*-th and *h*-th vectors given by

$$\cos\theta_{gh} = \frac{\sum_{i=1}^{N} X_{gi} X_{hi}}{\| X_g \| . \| X_h \|}$$

From (4),

$S_T^2 = S_X^2 - S_E^2$ and

$$\gamma tt(Theoretical) = \frac{S_T^2}{S_X^2} = 1 - \frac{S_E^2}{S_X^2} = 1 - \frac{\frac{1}{N}[\|X_g\|^2 + \|X_h\|^2 - 2\|X_g\|\|X_h\|\cos\theta_{gh}]}{NS_X^2}$$

[5]

Normally distributed *P*-scores help to know whether two subtests *g*-th and *h*-th are parallel if and only if $H_0 : \overline{p_g} = \overline{p_h}$ $H_0 : \sigma^2 p_g = \sigma^2 p_h$ and $\sigma^2{}_{p_g} = \sigma_{p_h}{}^2$ are accepted.

Equation (5), in terms of *P*-scores helps to test $H_0 : \gamma_{tt(Thepretical)=1}$ which boils down to test $H_0 : \sigma_x^2 = \sigma_T^2$ by *F*-test [32].

Discussion and Conclusion

Assumptions of standard parametric statistical methods are usually not satisfied by drug addiction data. Innovative use of statistical methods helps to better analyze and interpret substance abuse data, including better assessment of resultant QoL of substance users. Frequency based approaches [33] are simple to comprehend and calculate and involve no assumptions. 2 × *2* Contingency table can be extended to *r* × *S*contingency table for *r*-rows and *s*-columns to accommodate levels of drug addictions say mild, moderate, severe, etc. and various level of income or age-group. However, very low frequencies in off-diagonal cells can distort the result.

For only two outcomes [34] and a set of continuous independent variables, point bi-serial correlation is suitable and logistic regression model helps to assess and test significant logistic regression coefficient. Multilevel models are [35] encouraging to investigate effects at the cross-levels [36] (say, school effect on individuals) or third order cross-levels like *School*×*Individua* × *Classroom* etc. However, causal inference without randomization of all levels and cross-levels is a weak point since intercepts/slopes at higher levels depends on same at each lower level. But there are situations where randomization is not possible or may be unethical.

Proposed method of transforming ordinal raw scores of QoL items to follow normal distribution enables statistical analysis under parametric set up leading to meaningful comparisons and inferences, finding equivalent scores of [37]. two or more scales along with better computation of psychometric qualities like reliability, validity, responsiveness, etc. Considering theoretical advantages, the proposed method of transforming raw scores of QoL items to follow normal distribution is recommended. Future studies with longitudinal data can be undertaken to find sensitivity of the proposed score over time with emphasis on progression of disease and to different therapeutic interventions, etc.

Reference

1. United Nations Office on Drugs and Crime (2008).Treatment*:* International network of drug dependence treatment and rehabilitation resource centers Vienna, UNODC.
2. Gorski T, Miller M(1986).*Staying* Sober:A Guide for Relapse Prevention. Independence, MO: Independence Press.
3. https://link.springer.com/referenceworkentry/10.1007/978-0-387-78665-0_214
4. Bezarede Mekonnen and Legesse Alamrie (2020).Statistical Analysis of Risk Factors of Drug Addiction of Youth in Debre Berhan Town, Ethiopia. American Journal of Theoretical and Applied Statistics, 9 :1 (8-13) doi: 10.11648/j.ajtas.20200901.12
5. Gowing LR, Ali RL, Allsop S, Marsden J, Turf EE,(2015).Global statistics on addictive behaviours: status report. *Addiction*, 110:6 :904-19. doi: 10.1111/add.12899.
6. Peacock A, Leung J, Larney S, Colledge S, Hickman M, Et al(2018).Global statistics on alcohol, tobacco and illicit drug use: 2017 status report. Addiction; 113:10:1905-1926. doi: 10.1111/add.14234.
7. Yin RK, Kaftarian SJ, Yu P, Jansen MA(1997).Outcomes from CSAP's community partnership program:Findings from the national cross-site evaluation. Evaluation and Program Planning; 20:345–355.
8. Murray DM, Rooney BL, Hannan PJ, Peterson AV, Ary DV.(1994).Intraclass correlation among common measures of adolescent smoking: Estimates, correlates, and applications in smoking prevention studies. American Journal of Epidemiology; 140:1038–1050.
9. Krull JL, MacKinnon DP (2001).Multilevel modeling of individual and group level mediated effects. Multivariate Behavioral Research; 36: 249–277.
10. Cook TD,Campbell DT(1979).Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally
11. Livert D, Rindskopf D, Saxe L Stirratt M(2001).Using Multilevel Modeling in the Evaluation of Community-Based Treatment Programs. Multivariate Behavioral Research; 36 :2 155-184 10.1207/S15327906MBR3602_02
12. Smith KW, Larson MJ(2003).Quality of life assessments by adult substance abusers receiving publicly funded treatment in Massachusetts. The American Journal of Drug and Alcohol Abuse; 29, 323–335.
13. Moos R, Finney J, Cronkite R (1990) Alcoholism treatment: Context, process, and outcome. New York: Oxford University Press
14. https://pubmed.ncbi.nlm.nih.gov/10397415/.
15. Giacomuzzi SM, Riemer Y, Ertl M, Kemmler G, Rössler H Et al(2003).Buprenorphine versus methadone maintenance treatment in an ambulant setting: a health-related quality of life assessment. Addiction*,* 98:693-702
16. McHugo GJ, Drake R.E, Teague GB, Xie H(1999).Fidelity to assertive community treatment and client outcomes in the New Hampshire dual disorders study. Psychiatric Services; 50: 818–824.
17. Laudet AB, Morgen K White WL(2006).The role of social supports, spirituality, religiousness, life meaning and affiliation with 12-step fellowships in quality of life satisfaction among individuals in recovery from alcohol and drug problems. Alcoholism Treatment Quarterly; 24:33–73.
18. Laudet AB(2011).The case for considering quality of life in addiction research and clinical practice. Addiction Science & Clinical Practice; 6:44–55.
19. Morales-Manrique CC, Valderrama-Zurián JC, Castellano-Gómez M, Aleixandre-Benavent R, Palepu A;(2007).Cross cultural adaptation of the Injection Drug User Quality Of Life Scale (IDUQOL) in Spanish drug dependent population, with or without injectable consumption: Drug User Quality of Life Scale-Spanish (DUQOL-Spanish). Addict Behav;32:9:1913-21. doi: 10.1016/j.addbeh.2006.12.006.
20. Zubaran C, Emerson J, Sud R, (2012).The application of the drug user quality of life scale (DUQOL) in Australia. Health Qual Life Outcomes,2012; 10, 31 https://doi.org/10.1186/1477-7525-10-312012

21. Chung-g Lin, Jing-Shiang Hwang, Wen-Chung Wang, Wu-Wei Lai,(2019). Psychometric evaluation of the WHOQOL-BREF, Taiwan version, across five kinds of Taiwanese cancer survivors: Rasch analysis and confirmatory factor analysis, Journal of the Formosan Medical Association,118:1, Part 2, 215-222. https://doi.org/10.1016/j.jfma.2018.03.018.

22. Preston, Carolyn C, Colman AM(2000).Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences, Acta Psychologica; 104:1-15.

23. Guermazi M, Allouch C, Yahia M, (2012).Translation in Arabic, adaptation and validation of the SF-36 Health Survey for use in Tunisia. Ann Phys Rehabil Med, 5: 388–403.

24. Anagnostopoulos F, Niakas D, Pappa E(2005).Construct validation of the Greek SF-36 Health Survey. *Qual Life Res.*14:1959–1965

25. Lins L, Carvalho FM(2016).SF-36 total score as a single measure of health-related quality of life: Scoping review. *SAGE Open Medicine*; 4. 4:2050312116671725. doi: 10.1177/2050312116671725

26. Lewin RJP, Thompson DR, Martin CR, Stuckey N, Devlen J Eta al(2002).Validation of the Cardiovascular Limitations and Symptoms Profile (CLASP) in chronic stable angina. J Cardiopulm Rehabi. 22:184-191

27. https://link.springer.com/article/10.1007/s11205-007-9171-x

28. Chakrabartty Satyendra Nath (2020).Better Use of Scales as Measuring Instruments in Mental Disorders. Journal of Neurology Research Reviews & Reports, SRC/JNRRR-128, 1 – 7. DOI: doi.org/10.47363/JNRRR/2020 (2)117

29. Chakrabartty Satyendra Nath (2021).Integration of various scales for Measurement of Insomnia. Research Methods in Medicine & Health Sciences 2:3(102-111) DOI: 10.1177/26320843211010044

30. Parkerson, HA, Noel M, Gabrielle MP, Fuss, Samantha, Et al.(2013).Factorial Validity of the English-Language Version of the Pain Catastrophizing Scale–Child Version, *The Journal of Pain,* 14 :11, (1383-1389) https://doi.org/10.1016/j.jpain.2013.06.004

31. Chakrabartty Satyendra Nath, (2020b). Reliability of test battery: Methodological Innovations, 1–8. DOI: 10.1177/2059799120918340

32. United Nations Office on Drugs and Crime (2014) Treatnet: International network of drug dependence treatment and rehabilitation resource centres Vienna, UNODC.

33. Little TD, Schnabel KU, Baumert J (2000).Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples Mahwah, NJ: Erlbaum

34. Bow-Thomas CC,Velligan DI, Miller AL and Olsen Jerry (1999). Predicting quality of life from symptomatology in schizophrenia at exacerbation and stabilization, Psychiatry Research, 86 :2, 131-142, https://doi.org/10.1016/S0165-1781(99)00023-2

35. Arefnasab Z, Ghanei M, Noorbala AA, (2013). Effect of mindfulness based stress reduction on quality of life (SF-36) and spirometry parameters, in chemically pulmonary injured veterans. Iran J Public Health; 42: 1026–1033.

36. Boccard SG, Pereira EA, Moir L(2014).Deep brain stimulation of the anterior cingulate cortex: targeting the affective component of chronic pain. Neuroreport; 22: 83–88

37. Boi R, Racca L, Cavallero A,(2012).Hearing loss and depressive symptoms in elderly patients. Geriatr Gerontol Int; 12: 440–445