Comprehensive National Nutrition Survey

2016-18

Data Note

CNNS data version 1.0 dated 30.12.2019

Submitted to MoHFW

1. Background

The Comprehensive National Nutrition Survey (CNNS) is a nationally representative cross-sectional household survey collecting a wide range of data from more than 110,000 children and adolescents in both urban and rural areas in all 30 states of India. The target group includes preschoolers (children aged 0-4 years), school-age children (children aged 5-9 years), and adolescents aged 10-19 years. The CNNS was implemented by Ministry of Health and Family Welfare (MoHFW). The detailed anthropometric measurements collected from 112,316 children and adolescents and biological samples (blood, urine, and stool) from 51,029 children and adolescents from 2,035 primary sampling units (PSUs) across the country. The CNNS adopted a multi-stage sampling design with probability proportional to population size (PPS) after geographical stratification in rural and urban areas in all the states.

1.1 Data collection

The CNNS was conducted between February 2016 and October 2018. In CNNS, data were collected from 112,316 out of 116,371 sampled children/adolescents, yielding a response rate of 97%. Out of the 112,316 children 106,446 were assessed for anthropometric measurement yielding a response rate of 95%. Biological samples (blood, urine, and stool) from 51,029 children/adolescents out of 70,216 were collected and successfully analyzed, yielding a response rate of 73%.

2. Description of information collected in CNNS

The type of information in CNNS follows three broad domains: survey data, anthropometric measurements, and biomarkers.

2.1 Survey

Data have been collected by the research investigators through a face-to-face interview process to capture the behavioural information of eligible population by administering the different types of questionnaires.

2.2 Information from the household questionnaire

The household questionnaire listed all the usual residents in each household and any visitors who may have stayed at the house the night before the interview. For each person listed, information has been collected on age, sex, marital status and education. Also background information such as religion, caste of the household head and wealth quantities are included.

2.2.1 Information from the individual questionnaire

The individual questionnaire has collected the information on the selected eligible children and adolescents aged 0-19 years, who were usual residents of the sample household. Separate individual questionnaires were used for children and adolescents:
1. children aged 0-4 years,
2. children aged 5-9 years,
3. adolescents aged 10-14 years and
4. adolescents 15-19 years

The CNNS version 1.0 data has the background characteristics information- child/adolescent's date of birth, sex, parents 'education and exposure to mass media as given in CNNS chapter 3 and other chapters. Also the data has Information on IYCF practices from mothers and other caregivers of children aged 0-23 months using the World Health Organization (WHO) recommended definitions, methods, and questionnaire (WHO 2008). In older children and adolescents from 2-19 years, a seven-day frequency of consumption of the most important food items has been captured (as given in CNNS chapter 4).

2.3 Anthropometric measurement

After the interview was completed, seven types of anthropometric measurements were conducted (as given in CNNS chapter 5). All measurements except weight were taken two times following the standard procedures described by Gibson (Gibson 2000). In the analysis, all flagged measures were set to missing before calculating a mean of the double measure. The mean was used as the valid measure for calculation of the anthropometric z-scores.

2.4 Biomarkers

The CNNS provides following micronutrients, inflammatory and NCDs biomarkers for three age groups. The table provides the list of biochemical indicators and name of the corresponding variable in dataset (as given in CNNS chapters 6, 7 and 8).

Sr. No	Biochemical Indicator	Variable	Name of the corresponding variable
Data for children aged 1-4 years			
1	Anaemia	Altitude-adjusted Hemoglobin	Hgbn_alt_unic
2	Iron deficiency	Ferritin	Fer_unic
3	Folate deficiency	Folate	Rbcf_1unic
4	Vitamin B12 deficiency	Vitamin B12	Vtb12c_unic
5	Vitamin D deficiency	Vitamin D	Hydrov_unic
6	Vitamin A deficiency	Vitamin A	Vitaa_unic
7	Zinc deficiency	Zinc	Zns_unic
8	Urinary Iodine deficiency/ excess	Urinary Iodine	Uiod_unic
Data for children and adolescents aged 5-19 years			

1.	Anaemia	Altitude-adjusted Hemoglobin	Hgbn_alt_unic
2.	Iron deficiency	Ferritin	Fer_unic
3.	Folate deficiency	Folate	Rbcf_1unic
4.	Vitamin B12 deficiency	Vitamin B12	Vtb12c_unic
5.	Vitamin D deficiency	Vitamin D	Hydrov_unic
6.	Vitamin A deficiency	Vitamin A	Vitaa_unic
7.	Zinc deficiency	Zinc	Zns_unic
8.	Urinary Iodine deficiency/ excess	Urinary Iodine	Uiod_unic
9.	High Total cholesterol	Total cholesterol	Cho_dunic
10.	High LDL cholesterol	LDL cholesterol	Dirldl_dunic
11.	Low HDL cholesterol	HDL cholesterol	Chohdl_dunic
12.	High Triglycerides	Triglycerides	Tri_dunic
13.	High Plasma Glucose	Plasma Glucose	Glus_dunic
14.	High Glycosylated Hemoglobin	Glycosylated Hemoglobin	Hb_a1c_unic
15.	High Serum Creatinine	Serum Creatinine	Creat_unic

Note: For deficiency cut-offs, please refer the CNNS report

3. Content of data

There are three datasets for the three age group of children/adolescents based on information collected in CNNS.

Type of dataset	Sub-datafiles included			
	Household	Individual	Anthropometry	Biological
Children aged 0-4 years	Yes	Yes	Yes	Yes*
Children aged 5-9 years	Yes	Yes	Yes	Yes
Adolescents aged 10-19 years	Yes	Yes	Yes	Yes

*Biological samples were not collected from children less than one year

Key variables for merging data sets

State, PSU number, Household number, and respondent's line number are four key variables and are used to merge datasets. The identification variables below, are the same in all three datasets.

ID variables: Users may use either 'case_id' which is an unique number to identify a case in the file and based on combination of 'state', 'psu_no', 'hh_no', 'line_no'.
CNNS datasets are available in STATA (version 15/16). Datasets contain cases with complete individual interviews only. There are three data files named according to the age group of children/adolescents
1. CNNS_04_1.0.dta
2. CNNS_59_1.0.dta
3. CNNS_1019_1.0.dta

The sections below describes the content of each dataset, computed/recoded variables, survey weights and limitations in use of CNNS data to help users understand and use the data.

3.1 Nomenclature of variables

The name and numbering of variables has been given in such a way that users can easily identify variable. Although all the variables are well labelled, while using the dataset it will be helpful to have the questionnaire alongside.

- In all the three datasets, household-level variables correspond to the question number in the household questionnaire, preceded by 'h' or 'm' (for member information).
- Individual-level variables correspond to question number in the individual questionnaire and are preceded by 'q'.
- The suffix '_unic' is added in all the biomarker variables' name in the datasets.

3.2 Computed/recoded variables

Following key computed/recoded variables in all the three datasets are included for quick analysis. List of variable names and label are listed in table below:

Datafile: CNNS_04_1.0		
Recoded variables	Variable name	Label description
Household level	psu_no_ov	PSU(Unique within country)
	wi	Wealth Index for India
Individual level	agemons	Age of child in months
Anthropometry	muac	Mean MUAC
	height	Mean height
	tsft	Mean TSFT
	ssft	Mean SSFT
	clenhei	Converted length/height for deriving z score (in cms)
	cbmi	Calculated bmi=weight / squared(_clenhei)
	zwei	Weight-for-age z-score
	zlen	Length/height-for-age z-score
	zbmi	BMI-for-age z-score
	zwfl	Weight-for-length/height z-score
	zac	Arm circumference-for-age z-score
	zts	Triceps skinfold-for-age z-score
	zss	Subscapular skinfold-for-age z-score
	_fwfl	=1 if (_zwfl < -5 | _zwfl >5) (flagged cases)
	_flen	=1 if (_zlen < -6 | _zlen >6) (flagged cases)
	_fwei	=1 if (_zwei < -6 | _zwei >5) (flagged cases)
	_fbmi	=1 if (_zbmi < -5 | _zbmi >5) (flagged cases)
	_fhc	=1 if (_zhc < -5 | _zhc >5) (flagged cases)
	_fac	=1 if (_zac < -5 | _zac >5) (flagged cases)
	_fts	=1 if (_zts < -5 | _zts >5) (flagged cases)
	_fss	=1 if (_zss < -5 | _zss >5) (flagged cases)

Datafile: CNNS_59_1.0		
Recoded variables	**Variable name**	**Description**
Household level	psu_no_ov	PSU (Unique within country)
	wi	Wealth Index for India
Individual level	agemons	Age of child in months
Anthropometry	height	Mean height
	muac	Mean MUAC
	tsft	Mean TSFT
	ssft	Mean SSFT
	wc	Mean WC
	z_wc	Z-score Waist circumference
	z_tsft	Z-score TSFT
	z_ssft	Z-score SSFT
	z_muac	Z-score MUAC
	zwfa	Weight-for-age z-score
	zhfa	Height-for-age z-score
	zbfa	BMI-for-age z-score
	_fwfa	=1 if (_zwfa < -6 | _zwfa >5) (flagged cases)
	_fhfa	=1 if (_zhfa < -6 | _zhfa >6) (flagged cases)
	_fbfa	=1 if (_zbfa < -5 | _zfa >5) (flagged cases)

Datafile: CNNS_1019_1.0		
Recoded variables	**Variable name**	**Description**
Household level	psu_no_ov	PSU (Unique within country)
	wi	Wealth Index for India
Individual level	agemons	Age in months
Anthropometry	height	Mean height
	muac	Mean MUAC
	tsft	Mean TSFT
	ssft	Mean SSFT
	wc	Mean WC
	z_wc	Waist circumference Z-score
	z_tsft	TSFT Z-score
	z_ssft	SSFT Z-score
	z_muac	MUAC Z-score
	bp_sys1	Systolic-reading-1
	bp_dysys1	Diastolic-reading-1
	bp_sys2	Systolic-reading-2
	bp_dysys2	Diastolic-reading-2
	bp_sys3	Systolic-reading-3
	bp_dysys3	Diastolic-reading-3
	cbmi	Calculated bmi=weight / squared(height)
	zwfa	Weight-for-age z-score
	zhfa	Height-for-age z-score
	zbfa	BMI-for-age z-score
	_fwfa	=1 if (_zwfa < -6 | _zwfa >5) (flagged cases)
	_fhfa	=1 if (_zhfa < -6 | _zhfa >6) (flagged cases)
	_fbfa	=1 if (_zbfa < -5 | _zfa >5) (flagged cases)

3.3 Description of sample weights

Sample weights have been calculated for different level (national and states) and domain (survey/anthropometry and biological) of analysis and are described below. Please refer the CNNS report for further details. A cluster identification variable '**psu_no_ov**' is provided in the dataset for using the sample weight variables.

Type of data	Unit of analysis	Sample weight variable to be used
Survey and Anthropometry	National level	nat_weight_survey
	State level	state_weight_survey
Biological	National level	nat_weight_bio
	State level	state_weight_bio

The above sample weights are not appropriate to use with combined samples of different categories (i.e. combining different age groups or datasets). In case users are interested in combining age groups, appropriate pooled sample weights must be computed and used.

3.4 Limitations in use of CNNS data

Following are the limitations of CNNS data to use.

- CNNS is a cross sectional survey. The data cannot be used to conduct analyses of causality. It provides information on the associations between indicators and outcomes.
- Household level information was collected only to provide children and adolescents background characteristics. These data are not appropriate for the calculation of household level estimates.
- CNNS sample was designed to present results at state level. Analysis below the state level will not have sufficient statistical power to make any valid conclusion.
- Disaggregated analysis, for example by socio-economic status of CNNS biochemical indicators at state level cannot be done due to limitations in sample size. However such analysis at national level will be valid.
- The timing of CNNS data collection varies by states. Indicators affected by seasonality should be compared and interpreted with caution.

3.5 Additional data collected in CNNS survey

When the government gives official approval, the further data **CNNS data version 2.0** including additional indicators can be released on the topics below:

- Government schemes and services (coverage) including Vaccinations
- Patient Health Questionnaire (PHQ) - diagnostic instrument for common mental health indicators
- Developmental Delays (ECD indicators) and School Readiness (ASQ)
- Anthropometry of Mother
- Empowerment questions on gender
- Slum/non-slum
- Tobacco and alcohol consumption and NCD of parents
- Food Security
- Stool data