

Comparative Analysis of Adult and Childhood BMI Trends in China (2012-2022)

Shubham Parab Avinash Sahu Hoang Nguyen Lai Eswarthai Reddy

2025-11-06

Executive Summary

This report examines BMI trends in China from 2012 to 2022, covering both adults and children, and applies the CRISP-DM framework to reveal key patterns of obesity and underweight that shape public health.

1. CRISP-DM Cycle 1: Understanding Obesity Trends

1.1 Business Understanding

The primary goal of this investigation is to understand the distinct dynamics of the obesity epidemic in China by age group. Given that rising obesity rates pose a significant long-term public health and economic burden, this analysis will be crucial in addressing this issue. Specifically, we will compare the rate of change in age-standardised obesity prevalence between adults and children/adolescents over the last decade (2012 to 2022).

Research Questions: - How has obesity prevalence changed in adults versus children over the past decade? - Which demographic groups show the most significant increases? - What are the implications for public health policy?

1.2 Data Understanding

The analytic dataset (`merged_bmi_data`) was meticulously prepared using the `munge/01-A.R` script. This involved combining data from two distinct NCD-RisC sources, a reputable global health research network, and harmonising multiple clinical definitions of malnutrition (both obesity and underweight/thinness) into a single longitudinal format.

Data Structure: - Time period: 2012 to 2022 - Geographic focus: China - Demographics: Four groups (Adult Women, Adult Men, Child Girls, Child Boys) - Metrics: Obesity and Underweight/Thinness prevalence with 95% uncertainty intervals

```
# Display data structure
cat("Dataset dimensions:", nrow(report_data), "rows ×", ncol(report_data), "columns\n\n")

## Dataset dimensions: 88 rows × 9 columns

cat("Available metrics:", unique(report_data$Metric_Category), "\n\n")

## Available metrics: Obesity Underweight/Thinness

# Display sample of the cleaned data structure
head(report_data, 10) %>%
  knitr::kable(caption = "Sample of merged BMI data showing both obesity and underweight metrics")
```

Table 1: Sample of merged BMI data showing both obesity and underweight metrics

Year	Country	Sex	Prevalence	UI_Lower	UI_Upper	Age_Group	Metric	Metric_Category
2012	China	Women	0.0483711	0.0453724	0.0516099	Adult	Obesity	Obesity
2013	China	Women	0.0507449	0.0475999	0.0542301	Adult	Obesity	Obesity
2014	China	Women	0.0532124	0.0497560	0.0568621	Adult	Obesity	Obesity
2015	China	Women	0.0558037	0.0519113	0.0598507	Adult	Obesity	Obesity
2016	China	Women	0.0585190	0.0540091	0.0632275	Adult	Obesity	Obesity
2017	China	Women	0.0613450	0.0560180	0.0669262	Adult	Obesity	Obesity
2018	China	Women	0.0643024	0.0579682	0.0711004	Adult	Obesity	Obesity
2019	China	Women	0.0674118	0.0598056	0.0756880	Adult	Obesity	Obesity
2020	China	Women	0.0706985	0.0613815	0.0806926	Adult	Obesity	Obesity
2021	China	Women	0.0741682	0.0629796	0.0863296	Adult	Obesity	Obesity

1.3 Data Preparation

Data preparation was handled in the preprocessing script (`munge/01-A.R`), which: 1. Loaded adult and child BMI data from separate Excel files 2. Filtered for China and years 2012-2022 3. Created separate datasets for obesity and underweight/thinness metrics 4. Standardised column names across all datasets 5. Combined all data into a single tidy format with metric categories

1.4 Modelling: Calculating Absolute Change in Obesity

For our first analysis, we calculate the absolute percentage point (pp) change in obesity prevalence for each age-sex group between 2012 and 2022.

```
absolute_change_summary <- report_data %>%
  filter(Metric_Category == "Obesity") %>%
  group_by(Age_Group, Sex) %>%
  filter(Year %in% c(start_year, end_year)) %>%
  summarise(
    Prevalence_Start = Prevalence[Year == start_year],
    Prevalence_End = Prevalence[Year == end_year],
    Absolute_Change_pp = (Prevalence_End - Prevalence_Start) * 100,
    .groups = 'drop'
  )

# Format the data and then display the results
cat("\n**Table 1: Absolute Change in Obesity Prevalence in China (", start_year, "to", end_year, ")**\n")
```

Table 1: Absolute Change in Obesity Prevalence in China (2012 to 2022)

```
absolute_change_summary %>%
  mutate(
    `Prevalence Start (%)` = round(Prevalence_Start * 100, 1),
    `Prevalence End (%)` = round(Prevalence_End * 100, 1),
    `Absolute Change (pp)` = round(Absolute_Change_pp, 1)
  ) %>%
  select(Age_Group, Sex, `Prevalence Start (%)`, `Prevalence End (%)`, `Absolute Change (pp)`) %>%
  knitr::kable()
```

Age_Group	Sex	Prevalence Start (%)	Prevalence End (%)	Absolute Change (pp)
Adult	Men	4.4	8.9	4.6
Adult	Women	4.8	7.8	2.9
Child	Boys	9.0	15.2	6.2
Child	Girls	3.7	7.7	4.0

1.5 Evaluation (Cycle 1)

The initial analysis reveals several key findings:

- **Highest prevalence:** Adult populations show higher obesity rates than children
- **Largest absolute increase:** Child Boys with a 6.2 percentage point increase
- **Limitation:** Absolute change does not account for baseline differences or rate of change over time

Next Steps: A longitudinal visualisation is not only beneficial but also necessary to understand temporal trends and compare rates of change across groups.

2. CRISP-DM Cycle 2: Comprehensive Malnutrition Analysis

2.1 Business Understanding (Refined)

While Cycle 1 focused on obesity, a complete understanding of nutritional health requires examining both extremes of malnutrition. This cycle is a crucial step in expanding the analysis to include underweight (adults) and thinness (children), providing a holistic view of BMI trends in China.

2.2 Modeling: Longitudinal Visualization of Obesity Trends

```
# Prepare data for obesity
plot_data_obesity <- report_data %>%
  filter(Metric_Category == "Obesity") %>%
  mutate(Group = paste(Age_Group, Sex, sep = " - "))

longitudinal_trend_plot <- ggplot(plot_data_obesity,
  aes(x = Year, y = Prevalence,
    color = Group, fill = Group)) +
  geom_ribbon(aes(ymin = UI_Lower, ymax = UI_Upper),
    alpha = 0.15, color = NA) +
  geom_line(linewidth = 1.2) +
  geom_point(size = 2) +
  labs(
    title = "Obesity Prevalence Trends in China: Adult vs. Child/Adolescent (2012-2022)",
    subtitle = "Obesity defined as BMI >= 30 kg/m2 for Adults and BMI > 2SD for Children/Adolescents",
    y = "Age-Standardized Prevalence",
    x = "Year"
  ) +
  scale_y_continuous(labels = scales::percent) +
  scale_x_continuous(breaks = seq(start_year, end_year, 2)) +
  theme(legend.position = "bottom", legend.title = element_blank())

print(longitudinal_trend_plot)
```

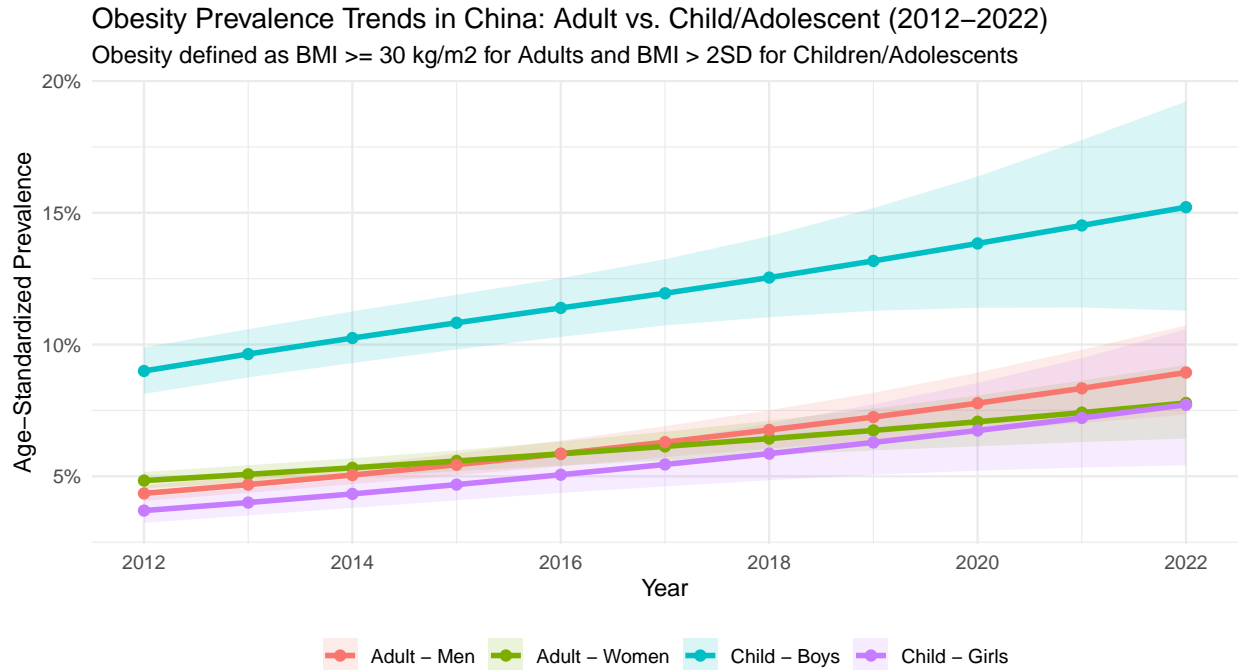


Figure 1: Longitudinal Trends in Obesity Prevalence (2012–2022) for Adults and Children/Adolescents in China, with 95% Uncertainty Intervals

2.3 Modeling: Underweight/Thinness Trends

To complete the malnutrition picture, we examine the prevalence of underweight in adults (BMI < 18.5 kg/m²) and thinness in children (BMI $< -2SD$). This provides crucial context to the obesity epidemic.

```
# Prepare data for underweight viz.
plot_data_underweight <- report_data %>%
  filter(Metric_Category == "Underweight/Thinness") %>%
  mutate(Group = paste(Age_Group, Sex, sep = " - "))

# Plot using the data using ggplot
underweight_trend_plot <- ggplot(plot_data_underweight,
  aes(x = Year, y = Prevalence,
    color = Group, fill = Group)) +
  geom_ribbon(aes(ymin = UI_Lower, ymax = UI_Upper),
    alpha = 0.15, color = NA) +
  geom_line(linewidth = 1.2) +
  geom_point(size = 2) +
  labs(
    title = "Underweight/Thinness Prevalence Trends in China (2012–2022)",
    subtitle = "Underweight defined as BMI < 18.5 kg/m2 for Adults and Thinness as BMI < -2SD for Children",
    y = "Age-Standardized Prevalence",
    x = "Year"
  ) +
  scale_y_continuous(labels = scales::percent) +
  scale_x_continuous(breaks = seq(start_year, end_year, 2)) +
  theme(legend.position = "bottom", legend.title = element_blank())
```

```
print(underweight_trend_plot)
```

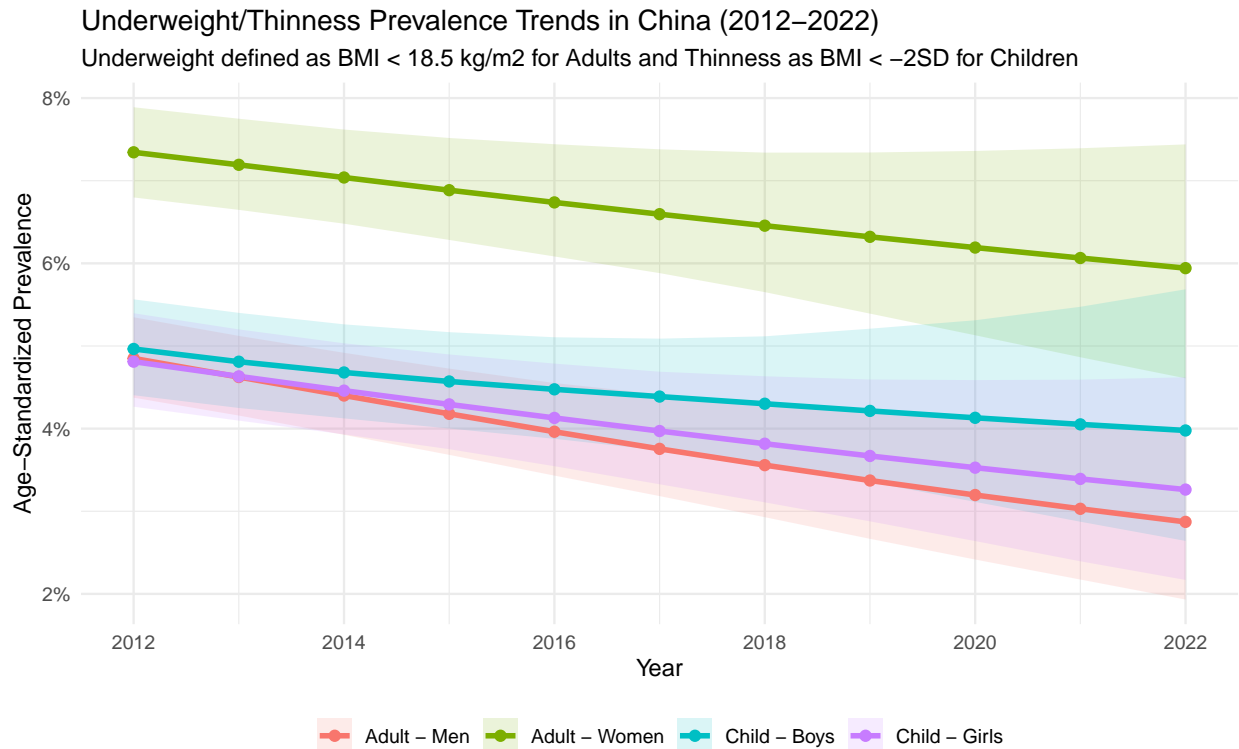


Figure 2: Underweight/Thinness Prevalence Trends (2012-2022) for Adults and Children/Adolescents in China, with 95% Uncertainty Intervals

2.4 Evaluation (Cycle 2)

The longitudinal chart shows valuable trends:

Obesity - Overall prevalence was increasing in all population groups - Adults had higher absolute prevalence than children, but the overall increase was similar - Uncertainty intervals were relatively small, suggesting estimates are robust.

Underweight/Thinness - Overall, underweight/thinness was generally declining for most groups - Declining trends were more evident for specific populations - Evidence of a "double burden" of malnutrition (eg, obesity increasing, persistent underweight)

2.5 Deployment

Key Takeaways and Policy Options:

1. Obesity Crisis on the Rise—Every population was increasing in obesity over 2012-2022; this is not just a trend, but an urgent crisis that requires immediate attention.
2. Double burden of malnutrition— This term refers to the simultaneous presence of both underweight/thinness and obesity in a population. In the context of China, it means that while obesity is increasing, there is also a persistent presence of underweight and/or thinness. This indicates that China is experiencing both ends of the malnutrition spectrum simultaneously.

3. Populations under consideration were cohort-specific for children and adults— differing trends over time do indicate the crucial need for age-targeted public interventions.

Recommendations: - Implement age-appropriate obesity prevention programs - Continue monitoring both obesity and underweight trends - Develop policies addressing the double burden of malnutrition - Invest in nutrition education and healthy food accessibility

3. Conclusion

This study revealed that an epidemic of rising obesity was accompanied by the persistent presence of underweight, as observed over 10 years in China (2012-2022). The study employed the CRISP-DM framework, a widely used data mining process model, to analyse the data. The findings illustrate the need for an integrated and age-appropriate public nutrition policy to address malnutrition.