

Lab Value Normalization

Unit Standardization

- **Glucose:** mg/dL ↔ mmol/L
- **Hemoglobin:** g/dL ↔ g/L
- **Creatinine:** mg/dL ↔ μmol/L
- SI units vs US conventional units

Reference Range Standardization

- Age-specific reference values
- Sex-specific reference values
- Pregnancy reference values
- Z-score calculation

Outlier Detection & Time Series Alignment

- IQR (Interquartile Range) method
- 3-sigma rule
- Time synchronization and missing value handling

1. Unit Conversion

Conversion Formulas for Key Lab Tests:

- **Glucose:** $\text{mg/dL} \times 0.0555 = \text{mmol/L}$
- **Hemoglobin:** $\text{g/dL} \times 10 = \text{g/L}$
- **Creatinine:** $\text{mg/dL} \times 88.4 = \mu\text{mol/L}$
- **Cholesterol:** $\text{mg/dL} \times 0.0259 = \text{mmol/L}$

When integrating multinational data, standardization to SI units is recommended.

2. Z-Score Normalization

$$Z = (X - \mu) / \sigma$$

X: measured value, μ : mean, σ : standard deviation

Application Examples:

- $Z > 2$ or $Z < -2$: Suspected outlier
- Apply age/sex-specific reference ranges
- Enable comparison across multiple lab tests

3. Outlier Detection Methods

IQR Method

Q1 = 25th percentile

Q3 = 75th percentile

IQR = Q3 - Q1

Outliers: $< Q1 - 1.5 \times IQR$ or $> Q3 + 1.5 \times IQR$

3-Sigma Rule

$\mu \pm \sigma$: Contains 68.3%

$\mu \pm 2\sigma$: Contains 95.4%

$\mu \pm 3\sigma$: Contains 99.7%

Outliers: $|X - \mu| > 3\sigma$

4. Time Series Data Processing

Missing Value Handling Strategies:

Forward Fill

Fill with previous value

Interpolation

Linear interpolation

Mean/Median

Replace with mean/median

For clinical data, expert review is recommended rather than simple imputation.

Normalization Process Visualization



💡 Practical Application Checklist

- ✓ Verify units by data source
- ✓ Preserve reference range metadata
- ✓ Check distribution before normalization
- ✓ Document conversion history
- ✓ Specify outlier handling criteria
- ✓ Ensure reversibility of conversions