

Introduction to Health Data Collection

Introduction to Health Data Collection

Key concepts, purposes, methods, and ethical considerations

1. What Is Health Data?

Category	Examples	Typical Sources
Clinical	Diagnoses, lab results, medication orders	Electronic Health Records (EHRs), hospital information systems
Administrative	Billing codes, admission/discharge dates	Claims databases, insurance registries
Patient? Generated	Symptom diaries, wearable sensor data, survey responses	Mobile apps, patient portals, wearables
Public?Health	Disease incidence, vaccination coverage, mortality rates	Surveillance systems, registries, census data
Research? Specific	Genomic sequences, biomarker levels, imaging data	Biobanks, clinical trial databases, research cohorts

2. Why Collect Health Data?

1. **Patient Care** – Enable accurate diagnosis, treatment planning, and continuity of care.
 2. **Quality Improvement** – Identify gaps, monitor performance indicators, and drive system? wide enhancements.
 3. **Public?Health Surveillance** – Detect outbreaks, monitor trends, and evaluate interventions.
 4. **Research & Innovation** – Support epidemiologic studies, develop predictive models, and test new therapies.
 5. **Policy & Planning** – Inform resource allocation, reimbursement policies, and health?system reforms.
-

3. Core Principles of Good Data Collection

Introduction to Health Data Collection

Principle	What It Means	Practical Tips
Relevance	Capture only data needed to answer the intended question.	Use a data dictionary; avoid “collect everything”.
Accuracy	Data must reflect the true value or event.	Standardized measurement tools; regular calibration.
Timeliness	Data should be available when needed.	Real-time feeds for critical alerts; batch uploads for non-urgent data.
Completeness	Minimal missing fields; all required variables present.	Mandatory fields with validation checks; follow-up for gaps.
Consistency	Uniform definitions and formats across sources.	Adopt coding standards (ICD-10, SNOMED CT, LOINC).
Security & Privacy	Protect confidentiality and comply with regulations.	Encryption, role-based access, audit trails.
Interoperability	Ability to exchange and interpret data across systems.	Use HL7 FHIR, CDA, or other open standards.

4. Common Data Collection Methods

Method	Setting	Strengths	Limitations
Electronic Health Records (EHRs)	Hospitals, clinics	Rich clinical detail; automated capture	Variable data quality; integration challenges
Claims & Billing Systems	Payers, insurers	Large population coverage; standardized codes	Limited clinical nuance; lag time
Surveys & Questionnaires	Community, research studies	Direct patient perspective; flexible	Recall bias; response fatigue
Wearables & Mobile Sensors	Home, remote monitoring	Continuous, high-frequency data	Data overload; device compliance
Registries	Disease-specific programs	Focused, longitudinal data	May miss non-registered patients
Laboratory Information Systems (LIS)	Labs, pathology	Precise quantitative results	Often siloed; requires mapping

Introduction to Health Data Collection

Method	Setting	Strengths	Limitations
Public?Health Reporting Systems	Health departments	Population?level trends	Under?reporting; reporting delays

5. Data Flow Overview (Typical Lifecycle)

1. **Planning** – Define objectives, variables, and data standards.
2. **Capture** – Input via forms, interfaces, sensors, or automated feeds.