A TECHNOLOGICAL AND INTELLECTUAL HISTORY

History of Medical Diagnostics Observability and Debugging

A history of clinical methods and diagnostic technologies

Diagnostic evidence and its interpretation

Concept and direction by Dmitry Vostokov, DumpAnalysis.org

Developed with AI assistance (GPT-6 Astra Max and Fable 5.1 Extra) and checked against the sources listed in this article.

Version 7 | September 2026

Historical synthesis with 150 references, chronology, glossary, and evidence taxonomy

Abstract

We can consider medical diagnostics as the recognition and interpretation of signs of human structure and behavior. Such signs include a patient's account of illness, findings from examination, changes in a specimen, physiological traces, and medical images. Their interpretation may lead to a diagnosis, suggest a possible mechanism, or show what additional evidence is needed. This history follows the development of these diagnostic activities and the instruments, institutions, and methods that made them possible.

The scope includes bedside examination, anatomy, laboratory medicine, imaging, electrophysiology, molecular and genomic testing, continuous monitoring, population surveillance, and computational diagnosis. We distinguish the acquisition of diagnostic evidence from its analysis, and both from the subsequent explanation of a patient's condition. For example, identifying an organism in a specimen and establishing that it caused the illness are different analytical steps. A similar distinction applies to a visible lesion, an abnormal waveform, or a genomic variant.

We use medical observability to describe the capacity of an evidence system to support inferences about otherwise hidden physiological and pathological states. The term is a retrospective analytical perspective. Debugging, as used in this series, concerns the investigation and correction of diagnostic processes and the revision of explanations when evidence disagrees. These comparisons have limits: patients have experiences, preferences, and rights, and diagnostic interventions can themselves cause harm. The recurring historical problem is how to acquire sufficient evidence, interpret it in context, and connect the resulting explanation to appropriate care.

How to read this history

This history is organized around diagnostic practices and their evidence. We follow selected developments from early case histories to contemporary computational methods. Egyptian, Mesopotamian, South Asian, Chinese, Greek, and Islamic traditions contributed different descriptions of illness and different ways of interpreting signs. Their categories cannot simply be translated into present-day disease names. The later global spread of European hospital and laboratory medicine also involved trade, empire, professional institutions, and unequal access to resources. This is a selective history of diagnostic methods, not an exhaustive history of every medical tradition or specialty.

For each development, we consider what could be observed, how the observation was obtained, and what conclusions it supported. We also examine the conditions under which it could mislead. For example, a blood-pressure value depends on the cuff and measurement procedure. A culture depends on the specimen and growth conditions. A genomic finding depends on the reference sequence and interpretation rules. The diagnostic process includes these supporting conditions even when they are absent from the final report.

References in square brackets identify the sources, numbered in order of first citation. Historical papers, archival records, standards, institutional publications, and research studies serve different purposes and should be read accordingly. An invention date does not necessarily mark the beginning of clinical use, and a publication does not establish sole priority. The four case studies show how observations acquired at different times and by different people were combined into explanations. The final chapter relates these examples to recurring diagnostic principles and, with explicit limits, to our pattern-oriented perspective. Figure 1 provides an illustrative summary at the end of the article.

Neighboring practices and their primary questions

Practice	Primary question	Characteristic evidence
Clinical assessment	What does the patient report, and what can be elicited or observed now?	History, examination, functional status, context, clinician–patient interaction
Measurement and testing	What quantity, structure, organism, molecule, or response is present under defined conditions?	Vital signs, assays, cultures, waveforms, images, pathology, genomic findings
Screening	Who among an apparently unaffected population may warrant further assessment?	Questionnaires, risk models, imaging, cytology, biochemical or genetic tests
Monitoring	How is selected physiology or disease burden changing over time?	Trends, telemetry, repeated specimens, alarms, symptom trajectories
Diagnosis	Which condition or mechanism best explains the available evidence?	Integrated history, examination, tests, prior probability, response and follow-up
Staging and prognosis	How extensive is disease, and what course or outcome is plausible?	Anatomical extent, biomarkers, function, longitudinal cohorts, prediction models
Surveillance	What patterns require population-level awareness or action?	Case definitions, laboratory reports, registries, incidence, clusters, wastewater data
Autopsy and diagnostic review	What was missed, misclassified, or causally misunderstood?	Postmortem findings, discordant results, timelines, preserved specimens, peer review

Table 1. Diagnostic practices and the evidence used for their principal questions.

RECURRING DEVELOPMENTS

- Clinical observations became more structured and increasingly quantitative.
- Anatomical findings could be compared with images and signals acquired during life.
- Laboratory procedures incorporated calibration, quality control, and automation.
- Repeated observations extended analysis beyond a single encounter or specimen.
- Probability and decision thresholds supplemented categorical descriptions.
- Shared records and population comparisons connected evidence across institutions.
- Computational methods added new ways to recognize patterns and review interpretations.

Contents

Seven parts, 35 chapters, and three reference appendices

Table 2. Contents and opening pages.

PART I

Foundations of medical diagnostics

1 Vocabulary and historical method

Let us first distinguish a test result from a diagnosis. A test produces an observation under specified conditions. Diagnosis requires the interpretation of that observation together with the patient's history, examination, time course, and possible alternative explanations. Screening, monitoring, staging, and prognosis use some of the same evidence but have different purposes. Screening identifies people who may need further investigation. Monitoring follows change. Staging describes extent, and prognosis concerns a possible future course.

We can identify three overlapping lines of development. Clinical methods organize the patient's narrative and findings obtained through examination. Laboratory methods examine material removed from the body under controlled conditions. Instrumental methods convert physical and biological processes into signals and images, which increasingly require computational processing. These are analytical distinctions rather than separate periods. A contemporary investigation may use all three, and disagreement between them may be the reason to continue the investigation.

The meaning of an abnormal result also needs to be specified. A reference interval describes a selected population measured by a particular method. It is not identical to a clinical decision threshold. Age, pregnancy, physiology, environment, treatment, and specimen handling can affect interpretation. Population categories, including ancestry, also require scrutiny because they may conceal social or environmental differences. A result can be measured correctly and still be interpreted incorrectly if its reference population or clinical purpose is unsuitable.

Medical observability therefore concerns the relationship between the evidence available and the questions we can answer. More measurements do not necessarily resolve an ambiguity. Debugging the diagnostic process may require checking identification, repeating an acquisition, reviewing an interpretation, or finding an independent source of evidence. Such work is directed at methods and explanations. It does not imply that a patient can be reproduced, reset, or modified in the way that a software system sometimes can.

2 Early diagnostic traditions

The Edwin Smith Papyrus provides examples of an early diagnostic structure: examination of an injury, description of signs, a judgment about the condition, and an expected course. The surviving manuscript dates from approximately 1600 BCE and preserves material with an earlier history. Mesopotamian

diagnostic and prognostic texts also organized signs into recurrent combinations. Their reasoning included empirical observation, ritual, and theoretical associations. We should examine these combinations within their historical setting instead of assuming that each corresponds to a modern diagnosis.[1–3]

Hippocratic case histories preserve sequences of observations. Fever, sleep, appetite, respiration, excretions, pain, and mental state could be followed over successive days. The sequence helped the observer recognize a course of illness and make a prognosis. In this setting, a diagnostic account did not have to identify an anatomical lesion or microorganism. Its object could be the changing condition of the patient. These records also show why the time of an observation and its relationship to earlier observations are part of the evidence.[4]

Pulse examination developed through several medical traditions. Greek and Islamic writers described qualities of arterial movement.[5, 6] Chinese traditions related pulse findings to broader classifications of bodily processes.[5, 7] South Asian medical practice also incorporated pulse examination into diagnosis.[5] Uroscopy used changes in urine as another source of diagnostic signs.[4, 6] The same visible sign could receive different interpretations within different explanatory systems.

We can recognize a general diagnostic activity in these examples: acquire observations, compare them with remembered or recorded cases, and relate the resulting pattern to an expected course. However, recognizing this common activity does not establish the validity of every historical explanation. Observations could be repeatable while their interpretation remained tied to assumptions that were difficult to test. The distinction between a recurrent pattern and its proposed mechanism will appear throughout this history.

Islamic medicine also contributed to the preservation, criticism, and organization of clinical knowledge. Al-Razi's case-based work and Ibn Sina's Canon belong to this history of diagnostic reasoning and transmission. Their writings combined inherited learning with their own organization and assessment of medical knowledge. Translation and teaching carried such works across languages and institutions. We should therefore distinguish the history of an instrument from the history of the practices through which observations were recorded, compared, and taught.[8, 9]

3 Anatomy and clinicopathological correlation

Anatomical investigation provided a way to relate an observed sign to a structure inside the body. Vesalius's De humani corporis fabrica, published in 1543, used dissection and illustration to examine inherited anatomical descriptions. Dissection had earlier histories, but printed illustrations made particular findings available for comparison by a much larger community.[10] Harvey's account of circulation in 1628 combined anatomy with experiment and quantitative argument. Here the explanation concerned movement through an organized system rather than an isolated structure.[11]

Clinicopathological correlation extended this comparison across time. Findings recorded during life could be compared with lesions found after death. Morgagni's De Sedibus et Causis Morborum of 1761 organized clinical cases and postmortem observations around diseased organs. Nineteenth-century hospital medicine expanded the available case collections through repeated examination, autopsy, statistics, and teaching. The diagnostic interpretation could then be revised when the anatomical evidence failed to match the original account.[12]

Cellular pathology introduced another level of analysis. Virchow related disease to changes in cells and tissues, supplementing the earlier focus on organs.[13] Each level also had its own sampling limitations. A biopsy contains selected tissue. A section preserves a thin plane through it. An autopsy examines the body after illness, intervention, and death have altered its condition. We cannot assume that a finding at one level gives a complete explanation of the whole process.

Autopsy subsequently became a means of reviewing the diagnostic process itself. A discrepancy could concern an unrecognized disease, an incorrect interpretation, or a failure to connect evidence already available. Postmortem findings are also subject to limitations, including sampling and changes after death. Their value as an audit comes from providing another evidence path against which the clinical explanation can be examined.[12, 14]

4 Clinical examination and assessment

The bedside examination combines information acquired in different ways. History establishes the sequence of symptoms, exposures, changes in function, and the patient's understanding of the problem. Inspection and palpation add visible and tactile signs. Auenbrugger's percussion, published in 1761, used differences in chest resonance to infer hidden conditions. Laennec's stethoscope, developed in 1816 and described in 1819, made internal sounds easier to localize, compare, and teach.[15–19]

These methods illustrate indirect diagnostic analysis. Dullness on percussion may be associated with fluid or consolidation. A murmur indicates a pattern of sound that must be related to blood flow, timing, and location. The observed pattern does not by itself identify a unique cause. Interpretation requires supporting information from the history and examination, and may require subsequent imaging or pathology. Instrument use also changes the encounter by directing attention toward some signs and away from others.

Case presentations, textbooks, and ward teaching made examination procedures reproducible across practitioners. Eponyms provided memorable names, although they could obscure mechanisms or contributions by other observers. Findings described as classic were not necessarily present in every stage of a disease or every patient population. Laboratory and imaging methods changed the role of examination without removing it. Examination may establish an initial probability, identify urgency, guide the next test, or reveal that a result does not fit the patient.

The observer is part of the acquisition process. Training, expectation, hearing, vision, touch, and the surrounding environment can affect a finding. Differences between observers therefore need examination just as differences between instruments do. The bedside remains a source of information about the person as a whole, but its findings benefit from comparison, documentation, and independent confirmation when appropriate.[15, 18, 20]

Psychiatric and cognitive assessment make the role of history and behavior particularly clear. The evidence may include reported experiences, speech, affect, attention, memory, behavior, daily functioning, and observations by other people. Operational diagnostic criteria and structured assessments make descriptions more comparable. They do not necessarily identify a single biological mechanism. WHO's 2024 clinical descriptions and diagnostic requirements for ICD-11 provide a contemporary example of systematic classification based on clinical findings.[21]

We need to distinguish reliability of classification from validity of explanation. Two practitioners may agree on a diagnostic category while important differences remain among the people assigned to it. NIMH's Research Domain Criteria initiative, begun in 2009, explores dimensions of behavior and biology across conventional categories. It is a research framework, not a replacement diagnostic manual. This distinction prevents a history centered on instruments from treating clinically organized narrative and behavioral evidence as merely incomplete laboratory diagnosis.[22]

5 Vital signs and functional testing

Quantified vital signs made observations easier to record and compare over time. Early thermometric work, including that associated with Santorio, was followed by instruments suitable for routine clinical measurement. Wunderlich's nineteenth-century temperature observations helped establish the fever chart. However, a temperature value depends on measurement site, method, time of day, and individual circumstances. The familiar value of 98.6°F is not a universal boundary separating health from illness.[23]

Counting the pulse and respiratory rate added explicit time intervals to older observations. Sphygmographs produced pulse traces, while the cuff and auscultatory methods associated with Riva-Rocci and Korotkoff made indirect blood-pressure measurement practical. The apparent simplicity of two reported numbers hides the measurement procedure. Cuff size, arm position, deflation, observer hearing, and repetition affect the result. Peripheral measurements also require interpretation in relation to the wider circulation.[24, 25]

The resulting numbers supported reference ranges, population comparisons, and treatment thresholds. They also made it possible to confuse numerical precision with diagnostic certainty. A single observation may be unrepresentative even when the instrument works correctly. A person's change from an earlier baseline can matter before a population threshold is crossed. Conversely, a difference between two measurements may come from a changed method rather than a changed physiological state.

These practices established a recurring acquisition process: observe under a defined procedure, preserve the time, compare the result, and decide whether action is needed. Later monitors automated parts of this process. Automation did not remove the need to check that the measurement belonged to the right person, represented the intended quantity, and reached someone able to interpret it.

Functional testing asks how a physiological system performs under specified conditions. Hutchinson's 1846 work on the spirometer related measured lung capacity to bodily characteristics and disease.[26] Later spirometry distinguished the amount of air exhaled from the speed of exhalation. Acceptability of the maneuver, repeatability, calibration, and reference equations became necessary parts of interpretation. The 2019 ATS/ERS technical statement documents how closely the test result depends on acquisition quality.[27]

This introduces a further distinction between observing a resting state and observing a response. A finding may emerge only during a defined task or after a controlled intervention. Such tests require the stimulus, timing, baseline, and response to be recorded together. The induced response is evidence about function under those conditions; it cannot automatically be generalized to every activity or to a unique disease mechanism.

PART II

Laboratory medicine and the classification of disease

6 Microscopy and causal organisms

Microscopy extended the range of structures available for inspection. Seventeenth-century observations disclosed cells, small organisms, and tissue details beyond unaided vision. Improvements in lenses were only one part of the subsequent development. Illumination, fixation, sectioning, staining, and preparation determined what could be distinguished and compared. An apparent structure might arise from the specimen, the preparation, or the instrument. Microscopic analysis therefore required knowledge of how the image was produced.[28]

Cell theory and Virchow's cellular pathology connected tissue changes with disease at a smaller scale.[13, 29] Pasteur's experiments and Koch's laboratory methods strengthened causal explanations involving microorganisms. The procedures later summarized as Koch's postulates connected association, isolation, experimental disease, and recovery of the organism.[30, 31] Their application was limited by carriers without symptoms, organisms that could not be cultured, mixed infections, and differences in host susceptibility. A useful causal procedure was not a universal rule for every disease.[32]

Staining made selected differences visible. Gram staining partitioned bacteria by their staining behavior and became useful for rapid classification. Acid-fast and other stains provided additional distinctions. Culture added an amplification step by allowing viable organisms to grow under selected conditions. The resulting growth depended on medium, atmosphere, temperature, incubation, and prior treatment. A negative culture therefore described what was recovered under those conditions, not everything that might have been present.[33–36]

Laboratory evidence also introduced errors between the patient and the interpretation. Misidentification, contamination, delay, and collection from an unsuitable site could produce a convincing result for the wrong diagnostic question. Organisms might represent colonization rather than the cause of illness. The specimen's origin and handling became necessary supporting information. Without them, the laboratory could identify a pattern correctly while the clinical interpretation remained wrong.

7 Urinalysis and clinical chemistry

Urine was accessible long before laboratory medicine developed. Its color, quantity, and sediment could be observed repeatedly. Chemical reactions subsequently provided more specific evidence about constituents such as glucose and protein. Microscopy added cells, casts, crystals, and organisms. We can

therefore regard urinalysis as a combination of evidence types rather than a single observation. Renal function, metabolism, hydration, collection, and storage all contribute to what is found.

Clinical chemistry converted many reactions into quantitative measurements through colorimetry, photometry, electrochemistry, enzymatic methods, and separation techniques. Glucose, electrolytes, enzymes, hormones, and other substances could be compared with reference materials and populations. The process came to be divided into pre-analytical, analytical, and post-analytical phases. These concern the preparation and handling of the specimen, the measurement itself, and the reporting and use of the result. Failure in any phase can invalidate the diagnostic conclusion.[37]

Skeggs's continuous-flow analyzer and subsequent automated systems increased the number and consistency of measurements that a laboratory could perform.[38–40] Automation also concentrated dependencies in calibration, reagent lots, instrument maintenance, and software. A systematic error could affect many results before it was recognized. Internal controls, external comparison, and reference methods were therefore needed alongside greater throughput. The absence of an instrument alarm did not establish that the entire examination process was correct.

Quality management extends beyond the analytical instrument. WHO's 2011 laboratory handbook treats personnel, equipment, purchasing, process control, information, documents, assessment, and corrective activity as related parts of laboratory work.[41] This helps explain why quality control of a reaction is insufficient when a specimen has been misidentified or a critical result has not reached the responsible clinician. The object of quality review is the process by which diagnostic evidence is acquired and used.

Clinical chemistry also requires comparisons to preserve their measurement context. A change of assay, calibration, or reference interval can create an apparent change in a patient's trajectory. Hemolysis, contamination, and other interferences can produce a result that is precise but unsuitable for interpretation. We should therefore ask whether serial values measure the same quantity under sufficiently comparable conditions before treating their difference as evidence of disease progression.

A biomarker inherits these dependencies. A measurable substance may be associated with injury, risk, burden, or treatment response. Its usefulness depends on the population, sampling time, threshold, and intended decision. The observation that two groups have different average concentrations does not by itself show that testing an individual will improve care. Chapter 24 returns to this distinction between measurement, association, and clinical use.

8 Hematology and blood groups

Blood examination required methods for preserving the sample, separating components, counting cells, and recognizing their morphology. Counting chambers and stained films made erythrocytes, leukocytes, and platelets available for systematic comparison. Automated counters later used impedance, optical scattering, and other measurements to produce counts and differential results at much greater scale. The

reported count was accompanied increasingly by distributions and flags that indicated when further examination was needed.[42]

Morphological analysis connects a count with the appearance of its components. Cell size, staining, inclusions, maturation, and distribution may suggest particular processes, but the same pattern can have several causes. Bone-marrow aspiration and biopsy provide information about production at another site and scale. Treatment, transfusion, pregnancy, altitude, and inflammation alter the context. A count, a blood film, and a marrow specimen are related observations rather than interchangeable descriptions.

Landsteiner's ABO investigations established differences between blood samples that were essential to understanding transfusion incompatibility.[43] Antiglobulin testing later detected antibodies that did not produce visible agglutination under the initial test conditions.[44] Crossmatching illustrates a useful diagnostic distinction: knowing the properties of two objects separately does not necessarily establish the result of their interaction. The proposed donor–recipient combination also has to be examined within an appropriate testing process.

Automated and manual methods can reveal different aspects of the same sample. An analyzer may detect a distribution that requires review, while a film reveals an abnormal form or a preparation artifact concealed by a summary value. Delta checks compare a result with earlier results from the same person. These comparisons support investigation, but a discrepancy still requires an explanation. It may reflect biological change, intervention, specimen error, or a difference in measurement.

9 Clinical microbiology

Clinical microbiology involves several questions that are easily conflated. Is an organism present? What is it? Does it explain the illness? Which antimicrobial agents inhibit it under laboratory conditions? Direct microscopy, culture, antigen detection, and nucleic-acid testing answer different parts of this investigation. Culture provides evidence of growth under specified conditions. A nucleic-acid result may remain positive when viable organisms are no longer recoverable. Neither finding can be interpreted independently of the specimen and clinical question.[33–36]

Identification methods progressed from appearance and biochemical reactions to automated panels, mass spectrometry, and sequencing. Comparison with a reference collection remains central. The method has to decide whether the observed characteristics sufficiently match a known category. Databases and taxonomies change, and unfamiliar or poorly represented organisms may be misidentified. A reported name should therefore be understood together with the method, confidence, and limitations of the comparison that produced it.

Antimicrobial susceptibility testing examines an organism's growth in relation to a drug. Disk diffusion, dilution, and automated methods require standardized inocula, media, incubation, measurement, and interpretive breakpoints. The resulting category informs treatment but does not reproduce the conditions

of the patient. Drug delivery, tissue penetration, host response, and the state of the infection also matter. The laboratory result is evidence about a defined interaction, not a guarantee of clinical success.[36]

Errors often arise when information crosses a boundary. A contaminant may be interpreted as the pathogen, a significant isolate may be dismissed, or an urgent result may fail to reach the person responsible for care. Interpretation therefore includes the specimen's origin, collection time, quality, and treatment history. A microbiology report should make its limitations and communication requirements visible. Correct identification alone does not complete the diagnostic process.

10 Serology and immunodiagnostics

Immunodiagnostics uses antigen–antibody interactions to obtain evidence about infection, immune response, proteins, hormones, and tissue targets. Early complement-fixation and precipitation methods extended the range of laboratory observations, while also creating problems of cross-reaction and standardization. The timing of the response was particularly important. Antibodies might be absent early in an illness and remain detectable after the relevant exposure. A serological result therefore refers to a biological history as well as to the time of sampling.[45]

Yalow and Berson's radioimmunoassay measured small concentrations through competitive binding and a radioactive readout.[46] Enzyme-linked immunosorbent assays subsequently used enzyme reactions to reveal binding, making many assays easier to distribute and automate.[47] Increased analytical sensitivity introduced its own interpretive requirements. Background binding, specimen effects, calibration, and the frequency of the target condition could all affect the meaning of a positive result.

The hybridoma methods of Köhler and Milstein provided a means of producing monoclonal antibodies with defined specificity.[48] Such reagents supported more reproducible assays and contributed to pathology and infectious-disease testing. Fluorescent antibody methods, developed earlier, made it possible to relate a detected target to its location.[49] Later immunohistochemistry combined recognition of molecular targets with preserved tissue morphology. Location could add information that a concentration measured from a mixed specimen would lose.

We should specify what a positive immunoassay actually demonstrates. It may indicate antigen, an immune response, a cross-reaction, or a signal exceeding a selected threshold. Confirmatory procedures, paired samples, and tests directed at different targets provide ways to examine these alternatives. Repeating the same interference with the same method does not produce independent confirmation. The acquisition mechanism and the biological mechanism both remain relevant to interpretation.

11 Histopathology and cytology

Histopathology preserves selected tissue structure for analysis. Fixation, embedding, sectioning, and staining convert a specimen into views that can be examined and compared. Biopsy allows this investigation during life, whereas autopsy provides a different range and context of sampling.[14] Frozen sections can answer time-sensitive questions during procedures, but their speed comes with preparation artifacts and limitations in the questions that can be resolved.[50]

Cytology examines cells obtained without retaining all of the surrounding tissue architecture. Papanicolaou and Traut's work on vaginal smears contributed to cervical cancer detection and the subsequent development of screening.[51] The value of a screening program depends on adequate sampling, interpretation, recall, confirmation, and access to treatment. A slide is only one component of this process. NCI's account of cervical screening concerns the contemporary practice; the historical publication supplies the earlier evidence.[52]

Morphology was progressively combined with immunohistochemistry, flow cytometry, cytogenetics, and molecular testing. A tumor classification can therefore depend on architecture, cell appearance, proteins, chromosome changes, mutations, and clinical findings. These observations may agree, refine one another, or create a need for review. The available tissue is finite, so the selection and ordering of tests also affect what can be established. Integrated reporting connects the findings instead of leaving each as an isolated result.[14]

Digital pathology made slides available as navigable image data for remote review, measurement, and computational analysis.[53, 54] The acquisition process then included the scanner, focus, color, compression, display, and software. A missing region or focus failure could remove diagnostic information before interpretation began. Validation has to address the intended diagnostic use of the whole system. A digital image retains a relationship to the glass slide, which in turn represents only selected material from the patient.

PART III

Seeing and measuring inside the living body

12 Radiography and contrast imaging

Röntgen's discovery of X-rays in 1895 made internal structures available for examination without incision. Early applications included fractures, foreign bodies, and chest disease.[55, 56] The radiograph was a projection in which tissues overlapped. Differences in attenuation produced the image, while positioning and exposure determined which relationships could be recognized. A visible outline did not by itself explain the underlying process, and superimposed structures could conceal an abnormality.

Fluoroscopy added observation of movement and support for procedures. Contrast agents increased the visibility of vessels and hollow organs whose boundaries were otherwise difficult to distinguish. These methods introduced additional risks from radiation and contrast administration. The subsequent development of radiological protection reflected experience with patient and occupational exposure. Justification of the examination and optimization for its diagnostic purpose became essential parts of practice.[57]

Image quality has to be assessed in relation to the question. Spatial resolution, contrast, noise, motion, and dose interact. Improving one property can worsen another, and a visually cleaner image does not necessarily answer the clinical question more accurately. Collimation, filtration, detectors, and standardized views influence the evidence produced. A normal radiograph cannot exclude a process below its resolution or outside the region and projection examined.

Radiology also produced a durable diagnostic record that other observers could review. Images could be compared with earlier examinations, annotated, and eventually transmitted. Such comparison made changes visible, but it also introduced findings unrelated to the original question. These incidental findings require an interpretation of their own. Their presence does not automatically justify further intervention, although they may create an obligation to explain uncertainty and arrange appropriate follow-up.

13 Electrophysiology and hearing assessment

Einthoven's string galvanometer and standardized leads made cardiac electrical activity available as a trace.[58] The electrocardiogram represents potential differences measured from selected positions. Its waves and intervals describe an electrical process over time rather than a direct image of the heart. Interpretation connects this representation with the arrangement and behavior of the conduction system and myocardium.

Waveform morphology and timing provided evidence about rhythm, conduction, ischemia, electrolyte disturbances, and other conditions. Standardized lead placement supported comparison between recordings. However, misplaced electrodes, motion, filtering, poor contact, and pacing could change the pattern or imitate disease. The recording conditions therefore belong to the interpretation. A recognizable waveform pattern can suggest several mechanisms, and the patient's history and other findings help distinguish them.

Electroencephalography extended electrical recording to the brain, including Berger's human recordings and their publication in the 1920s.[59] Electromyography, nerve-conduction studies, and evoked responses examine other parts of the nervous system. Averaging and frequency analysis can help identify activity that is difficult to recognize in an individual segment. State, medication, stimulus, electrode placement, and artifact remain important. A trace has to be related to the behavior and conditions present when it was acquired.

Automated analysis was particularly attractive for waveforms because intervals, amplitudes, and recurring shapes could be measured computationally. The resulting report could assist comparison and triage. It could also become a source of diagnostic anchoring if the reader accepted its interpretation before checking acquisition quality and clinical context. This separates two tasks that automation may combine: measuring a feature and assigning meaning to it.

Hearing assessment adds an example in which the response can be measured through different physical processes. Behavioral tests depend on the person's response to sound, whereas auditory brainstem responses record electrical activity following stimulation. Kemp's 1978 description of stimulated acoustic emissions showed that sound generated within the auditory system could itself be measured.[60] Otoacoustic-emission and auditory-brainstem-response methods subsequently became important in newborn hearing screening. They sample different functions and require diagnostic follow-up when screening indicates a possible problem.[61]

14 Ultrasound and echocardiography

Ultrasound uses the transmission and return of sound to obtain information about tissue. Early obstetric and abdominal applications related echo timing and amplitude to internal structures, while echocardiography applied related methods to the moving heart.[62, 63] The method could be repeated without ionizing radiation and brought imaging closer to the bedside. Its usefulness depended on selecting an acoustic path through which the relevant structures could be examined.

Doppler methods provided evidence of motion through frequency shifts. Different implementations separated velocity, direction, and location in different ways. A continuous-wave measurement can detect high velocities but cannot assign each signal to a unique depth. Pulsed methods provide depth selection but introduce other limits, including aliasing. Angle, gain, pressure, and the acoustic window affect the

result. A displayed flow pattern is therefore a measurement made under particular geometric and instrument conditions.

The ultrasound image is constructed from assumptions about sound propagation, reflection, and attenuation. Shadowing, enhancement, reverberation, and other effects can either obscure a structure or help characterize it. Acquisition and interpretation are closely connected because the examiner chooses and adjusts the view while observing the result. A saved image preserves only part of that activity. The diagnostic record benefits from documenting which regions and views were adequately examined.

Portable systems extended ultrasound to wards, emergency care, ambulances, and settings without a conventional imaging department. A smaller instrument does not by itself establish a complete diagnostic service. Focused examinations require a defined question, appropriate training, records, quality review, and a path for further assessment when findings are inconclusive. The intended scope of the examination determines what a negative finding can reasonably exclude.

15 Nuclear medicine

Tracer methods examine the movement and distribution of a labeled substance. De Hevesy's work established radioactive tracers as a means of investigating biological processes.[64] Diagnostic nuclear medicine combined a radiopharmaceutical with detection equipment and a model of its expected behavior. The location of activity could then provide evidence about perfusion, metabolism, or another physiological process, depending on the tracer and acquisition time.

Anger's scintillation camera made planar gamma-ray imaging practical.[65] Kuhl and Edwards's section-scanning work was an early step toward emission tomography.[66] Rotating detectors and reconstruction subsequently supported SPECT, while PET used the detection of coincident annihilation photons to estimate the distribution of a positron-emitting tracer. Fluorodeoxyglucose became widely useful for examining altered glucose uptake in oncology and other fields. These developments added spatial information to physiological measurements and created new dependencies on detectors, timing, correction, and reconstruction.[67, 68]

Tracer uptake rarely identifies a unique mechanism. Malignancy, inflammation, repair, and normal activity may share aspects of a distribution. Preparation, blood glucose, uptake time, motion, and calibration affect quantitative measures. The interpretation therefore requires knowledge of what the tracer measures and the alternatives that can produce a similar pattern. A standardized uptake value is a derived measurement, not an independent diagnosis.

Hybrid PET/CT and SPECT/CT relate functional findings to anatomical structures. Combining modalities can narrow the possible explanations, but registration errors and different acquisition times can also create disagreement. We should preserve the source and processing history of each component.

Agreement is most informative when the components provide sufficiently independent evidence rather than repeat the same assumption in different displays.

16 Computed tomography

Computed tomography reconstructs cross-sectional attenuation from multiple X-ray measurements. Cormack's mathematical work and Hounsfield's scanner development made this approach clinically practical, and their contributions were recognized by the 1979 Nobel Prize.[69–71] CT reduced the superposition that limits projection radiography. Structures previously concealed behind one another could be examined in separate cross-sections.

The CT image depends on a computational acquisition and reconstruction process. Detector calibration, geometry, reconstruction kernel, slice thickness, contrast timing, and display window all affect what the observer sees. Hounsfield units relate attenuation to a reference scale, but partial-volume effects and acquisition conditions limit simple interpretation of a pixel value. The reconstructed image is a model-based representation of the measurements rather than an unprocessed view of tissue.

Faster acquisition, helical scanning, multidetector systems, and cardiac gating extended the range of practical examinations. CT became useful in trauma, vascular imaging, pulmonary assessment, and image-guided procedures. Wider use also increased attention to repeated exposure and incidental findings. The question is whether a selected acquisition provides the needed information at a justified dose, not whether the largest possible amount of image data can be collected.[57]

CT makes the distinction between source data and derived evidence particularly clear. Most readers examine reconstructed images rather than detector measurements. A different reconstruction can alter noise, texture, and the visibility of small features. Iterative and learned reconstruction methods add further processing choices. For later review, the protocol and software context help explain why an image looks as it does and whether it can be compared with an earlier examination.

17 Magnetic resonance and functional imaging

MRI uses nuclear magnetic resonance signals, magnetic-field gradients, and reconstruction to obtain spatial information. Lauterbur demonstrated spatial encoding, while Mansfield and Maudsley published line-scan NMR imaging.[72, 73] Mansfield's separate 1977 paper described a faster approach using spin echoes.[74] Lauterbur and Mansfield received the 2003 Nobel Prize for their contributions to MRI.[75] The method provides several forms of tissue contrast without ionizing radiation. Consequently, an MRI examination is a selected collection of measurements, each emphasizing particular properties rather than reproducing one natural appearance of the body.

Pulse sequence, field strength, coil, timing, motion, and reconstruction determine the resulting contrast. Relaxation, diffusion, susceptibility, and perfusion measurements provide different information. This flexibility also creates protocol dependence and artifacts. Safety includes the compatibility of implants and equipment with the magnetic environment, as well as heating, acoustic noise, and projectile risks. The absence of ionizing radiation does not make every acquisition free of risk.

Blood-oxygen-level-dependent functional MRI uses changes related to blood oxygenation as indirect evidence of neural activity.[76] Human activation studies followed in 1992.[77] A functional map depends on task design, acquisition, movement, processing, and statistical analysis. The relationship between neural activity and the vascular response is another part of the inference. Such a map should therefore be interpreted as a result of an experiment and model, with their limitations, rather than as a direct picture of a mental event.

We can distinguish several transformations in MRI: physical behavior produces a signal, the signal is sampled, reconstruction produces an image, and further processing may produce a map or measurement. Interpretation adds another level. Preserving acquisition metadata and the relationship between these derived objects supports later reanalysis. Otherwise, a change in the processing method can be confused with a change in the patient.

18 Endoscopy and ophthalmic imaging

Endoscopy brings illumination and an optical view into a body cavity. Early rigid instruments were limited by access, illumination, discomfort, and risk. Improved lenses and fiber-optic bundles supported flexible instruments for examining the digestive and respiratory tracts. Video systems subsequently made the view available to several observers and allowed it to be recorded.[78, 79]

The examination is an activity as well as an image. The operator chooses where to look and when to wash, distend, magnify, biopsy, or intervene. Surface appearance may suggest a finding that requires tissue examination at another scale. Inadequate preparation, an unexamined area, or a lesion behind a fold can prevent detection. A collection of normal-looking saved images does not necessarily establish that the intended region was completely examined.

Changes in illumination and optical processing can reveal additional tissue features. Narrow-band and fluorescence methods alter contrast, while confocal techniques and optical coherence tomography provide other forms of spatial information. Computer-aided detection can direct attention toward a possible lesion. Each method adds its own acquisition conditions and potential errors. Its contribution has to be evaluated within the procedure in which the observer receives and acts on the information.

Optical diagnosis also developed through examination of the eye. Helmholtz's 1851 ophthalmoscope allowed inspection of the living fundus and helped relate visual impairment to visible internal findings.[80] Subsequent retinal imaging made these observations easier to preserve and compare. This is a distinct

diagnostic development from endoscopy: light can reach relevant structures through the eye's own optical path, although the view still depends on illumination, alignment, and the transparency of intervening tissues.

Huang and colleagues' 1991 optical coherence tomography paper described cross-sectional imaging using low-coherence interferometry.[81] OCT added depth-resolved tissue information at a scale different from conventional fundus photography. An OCT scan, a visual-function measurement, and a reported symptom answer related but different questions. Their comparison can reveal discordance that any one examination would miss. The saved optical image should remain connected to the method, acquisition quality, and clinical purpose.

For endoscopy and other interactive examinations, review must also consider the acquisition path. Coverage, landmarks, preparation, complications, specimen identity, and follow-up are part of the evidence. The final interpretation should distinguish an adequately examined normal region from a region that was not accessible or not recorded. This is one reason why the diagnostic process cannot be assessed solely from selected images.

19 Digital imaging and remote interpretation

Digital imaging separated the image from photographic film and made it available for copying, processing, and transmission. PACS linked acquisition devices with storage, worklists, displays, and reporting. Teleradiology allowed acquisition and interpretation to take place at different locations. These arrangements extended access while introducing dependencies on network availability, patient identification, retrieval of prior studies, and communication of urgent findings.[82, 83]

ACR–NEMA work in the 1980s preceded the DICOM standard, which provided structures for imaging objects, metadata, and network communication.[82] The image could travel with information about the patient, study, series, and acquisition. Standardized transport did not remove differences in local terminology or workflow. Correct interpretation still depended on appropriate mappings, implementation, and preservation of meaning across systems.

Digital processing produced additional representations from the same acquisition. Windowing, multiplanar views, subtraction, segmentation, measurements, and computer-generated prompts could disclose useful patterns. They could also make it difficult to determine which object was original and how a derived result had been obtained. Source relationships and processing versions therefore became part of diagnostic provenance. Two images derived from one dataset should not automatically be counted as independent evidence.

Successful transmission is one stage in a longer process. The correct study must reach the correct record, be technically adequate for the question, and be interpreted with relevant earlier evidence. The report then has to reach the responsible clinician and patient, with follow-up where necessary. Availability and turnaround measurements are useful, but they do not by themselves establish that the diagnostic information affected care.

PART IV

Genetic molecular and point of care diagnosis

20 Cytogenetics and newborn screening

The 1956 demonstration by Tjio and Levan that the usual human chromosome number is 46 corrected a long-standing count and helped establish modern human cytogenetics.[84] Improved cell culture, chromosome preparation, staining, and later banding made numerical and structural differences easier to identify. The association between an additional chromosome 21 and Down syndrome connected a recognizable clinical condition with a chromosomal finding.[85] These developments depended on better acquisition as well as changed interpretation.

A karyotype represents selected cells prepared at a suitable stage of division. Tissue choice, culture, mosaicism, resolution, and interpretation affect what the examination can detect. Chromosomal microarrays subsequently provided finer evidence of copy-number differences, but did not replace every capability of a karyotype. For example, a balanced rearrangement may require another method. Increased resolution also produces findings whose clinical significance is uncertain.

Newborn screening combined specimen collection with a population program. Guthrie's bacterial inhibition test for phenylketonuria and the use of dried blood spots provided a practical way to collect and transport small samples. Subsequent programs added conditions and methods, including tandem mass spectrometry and molecular tests. The diagnostic activity moved beyond individual specialist referrals to an organized sequence of screening, confirmation, and early intervention.[86, 87]

The benefit depends on the complete sequence. Collection time, specimen quality, thresholds, repeat testing, confirmation, communication, and treatment availability all affect the result. A positive screen identifies a need for further investigation; it does not by itself establish every target diagnosis. Retained samples also create questions about consent, storage, and secondary use. The laboratory method and the organization of follow-up cannot be separated when assessing the program.

21 Molecular diagnostics

Molecular diagnostics made selected nucleic-acid sequences available as diagnostic evidence. Restriction enzymes produced fragments, electrophoresis separated them, and Southern's 1975 method transferred DNA to a membrane for detection with a probe.[88] These procedures supported the analysis of inherited variation and rearrangements. Their results depended on a selected target and an expected pattern, so an untested sequence could remain outside the investigation.

Fluorescence in-situ hybridization added spatial information by locating labeled probes within chromosomes, cells, or tissues.[89] The analyst could connect the number and arrangement of signals with a selected sequence or rearrangement. Hybridization quality, background, cell selection, and signal interpretation affected the result. FISH illustrates targeted observability: it can answer a defined question precisely while leaving many other genomic questions unanswered.

PCR increased the amount of a selected DNA target through repeated amplification.[90] Mullis's retrospective account dates its conception to 1983.[91] Early diagnostic work included the analysis of sickle-cell variants.[92] Reverse transcription allowed RNA to be examined through a DNA intermediate, and quantitative PCR related the evolving amplification signal to the initial target amount. Scarce material could then become detectable in a practical time. The target sequence and the conditions of amplification remained central to the assay.

Amplification also increases the consequences of contamination and other process errors. Inhibition, extraction loss, primer mismatch, and unsuitable sampling time can prevent detection. Carryover can produce a positive signal that does not originate from the intended specimen. Controls and separation of procedures help distinguish these possibilities. A cycle threshold should be interpreted within the assay and specimen context; it is not a universally comparable measure of disease or infectiousness.

Digital PCR changes the measurement by partitioning material and assessing individual reactions. Vogelstein and Kinzler's 1999 paper described detection and quantification of rare mutant sequences through this approach.[93] Counting positive and negative partitions can estimate target abundance, but the estimate still depends on sampling, partition volume, classification, and amplification performance. A numerical count does not remove the uncertainty introduced before the material reached the instrument.

CRISPR-based detection added programmable sequence recognition to the molecular repertoire. Gootenberg and colleagues' 2017 SHERLOCK study combined amplification with Cas13a activity and a reporter signal to detect selected nucleic acids.[94] Such work demonstrated additional ways of connecting a target sequence to a readable result. Demonstrating a detection mechanism in a study is distinct from validating a complete clinical assay for a specified population and setting.

22 Sequencing and genomic diagnosis

Sanger's chain-termination method provided a practical way to determine a DNA sequence.[95] Clinical interpretation then had to connect that sequence with inheritance, phenotype, population frequency, functional evidence, and alternative explanations. The sequence read was an acquired result; the claim that a variant explained a condition was a further inference. Better reading accuracy could resolve one source of uncertainty without resolving the other.

The Human Genome Project and parallel sequencing efforts produced reference assemblies and shared infrastructure for much broader comparison.[96–98] The completion announced in 2003 marked the

project's achievement, but did not mean that every repetitive region or every human genomic variant had been resolved. A reference is a coordinate and comparison system. It is not a universal definition of a healthy human genome, and its population representation affects subsequent analysis.

Next-generation sequencing increased throughput by performing large numbers of reactions in parallel.[99] Panels, exomes, genomes, tumor samples, pathogens, and RNA could be examined at new scales. Coverage, alignment, duplicated reads, capture bias, repeats, structural variation, mosaicism, and contamination became important quality dimensions. The analysis also depended on computational versions and reference data. A negative report needed to distinguish regions adequately examined from regions the method could not reliably assess.

ACMG/AMP guidance provided a structured framework for combining evidence when classifying sequence variants.[100] Classification can change as new family, functional, or population information becomes available. Laboratories therefore have to distinguish a new analysis of stored data from a new interpretation of an existing finding. Consent, secondary findings, recontact, and communication across generations extend the process beyond the original specimen and report.

Later reference work addressed some limitations of earlier assemblies. The 2022 T2T-CHM13 sequence resolved previously unfinished regions of the autosomes and X chromosome; that assembly did not include a Y chromosome.[101] The 2023 draft human pangenome used assemblies from multiple individuals to represent more variation than one linear reference could contain.[102] These were developments in the evidence infrastructure. They did not make every clinical genome complete or every detected variant interpretable. Reference version and analysis method remain necessary parts of a diagnostic report.

23 Prenatal diagnosis

Prenatal investigation combines evidence about fetal development with decisions made under considerable uncertainty. Ultrasound provides observations of anatomy, growth, gestational age, and placental relationships. Amniocentesis and chorionic-villus sampling provide material for chromosomal, biochemical, or molecular analysis. The procedures differ in timing, sampled tissue, limitations, and risk. A result needs to be interpreted in relation to the question that led to testing.[103]

The finding of cell-free fetal DNA in maternal plasma in 1997 opened another route to prenatal assessment.[104] Sequencing subsequently enabled counting approaches for selected chromosomal differences.[105] Much of the relevant material originates from the placenta. For common aneuploidy applications, these methods are screening tests whose results alter probability. The name of an early research paper using the word diagnosis should not be taken to mean that every subsequent cell-free-DNA result is independently diagnostic.

Placental mosaicism, maternal findings, low fetal fraction, vanished twins, gestational age, and technical factors can contribute to discordant results. Some results therefore require confirmation with a different specimen and method. The term non-invasive refers to obtaining maternal blood rather than performing an invasive fetal or placental procedure. It does not describe the full consequences of the information or the decisions that may follow.

Counseling should make the test's scope, possible results, uncertainty, and follow-up understandable. The pregnant person's preferences and reproductive autonomy are part of the process. Expanding technical capability without access to counseling, confirmation, and support can increase the burden of uncertain findings. Diagnostic progress must therefore be assessed in relation to the care available after the result.

24 Biomarkers and liquid biopsy

Biomarkers have different intended uses. The FDA–NIH BEST resource distinguishes, among others, susceptibility, diagnostic, monitoring, prognostic, predictive, response, and safety roles.[106] A marker useful for following established disease may perform poorly in population screening. A predictive marker concerns response to a particular intervention, whereas a prognostic marker concerns a future outcome in a defined context. The role has to be stated before performance can be assessed.

Tumor-associated markers such as PSA and CA-125 illustrate the limitations of interpreting a concentration as a disease name. Their values require clinical context and can be affected by processes other than the target malignancy. Cardiac troponin belongs to a different category: it is used as evidence of myocardial injury. The 2018 Universal Definition distinguishes myocardial injury from myocardial infarction, for which evidence of ischemia is also relevant.[107] These examples show why a biomarker's biological target and diagnostic use should be described separately.

Liquid biopsy examines circulating material, including tumor DNA or cells, obtained from blood or other fluids.[108] It can provide information about selected mutations, disease burden, resistance, or recurrence when those uses are validated. Low abundance, variable shedding, clonal hematopoiesis, sample handling, and analytical thresholds complicate interpretation. A negative result may reflect insufficient detectable material rather than absence of disease. A positive variant may also require evidence about its source.

We can distinguish three questions. Analytical validity concerns whether the method measures the intended target reliably. Clinical validity concerns the relationship between that measurement and a condition or outcome. Clinical utility concerns whether using the result improves decisions or outcomes. These questions are related, but evidence for one does not automatically answer the others. Serial changes also require attention to biological and analytical variation before they are interpreted as progression or response.

25 Point of care and home diagnostics

Point-of-care testing shortens the distance between measurement and a possible decision. Examples include urine strips, glucose meters, pregnancy tests, blood-gas instruments, rapid antigen tests, and cartridge-based molecular assays. Clark and Lyons's enzyme-electrode work provided an important basis for biosensing.[109] Moving a test closer to the patient changes transport and turnaround, but it also changes who acquires the specimen, operates the device, and interprets the result.

A lateral-flow device combines sample movement, recognition reagents, capture, and a visible or instrument-read signal.[110] Its simple outward form contains several dependencies. Sample amount, timing, temperature, storage, flow, and reagent quality affect performance. A control line demonstrates selected process conditions; it does not prove that the sample was taken correctly or that every possible cause of a false negative has been excluded.

Decentralized testing requires quality arrangements appropriate to its setting. Operator training, maintenance, controls, result identification, connectivity, and confirmatory procedures may be less visible than in a central laboratory. These functions still have to be provided. A fast result is useful when its reliability and meaning are understood and when someone can act on it. Turnaround alone is an incomplete measure of diagnostic performance.[111]

Home testing extends acquisition and initial interpretation to the person using the test. This can improve access, privacy, and repeated observation. It also makes instructions, language, accessibility, and follow-up part of the assay's practical use. A result may need explanation or confirmation, and an unreadable result should be represented as such. Testing is incomplete when the person receives a finding but has no workable route to further care.

Access to diagnostics also became an explicit health-system objective. WHO's first Model List of Essential In Vitro Diagnostics was introduced in 2018 and published in Technical Report Series 1017 in 2019.[112] It distinguished testing at different levels of care and supported national selection according to needs and resources. This adds a historical perspective often lost in accounts of invention: an established test may remain unavailable because supply, electricity, maintenance, transport, trained staff, or referral is missing. Availability of the complete service determines what can actually be diagnosed.

PART V

Continuous and population medical observability

26 Critical care monitoring

Critical care brought repeated observation into close contact with intervention. The response to the 1952 Copenhagen poliomyelitis epidemic, including Ibsen's use of positive-pressure ventilation and concentrated care, contributed to the development of modern intensive-care practice.[113] Later bedside systems combined ECG, pressure, temperature, and respiratory measurements. Their purpose was to make deterioration and response visible quickly enough for clinical action.

Blood-gas measurements, pulse oximetry, and capnography observe related but different physiological processes. Blood-gas electrodes quantify selected properties of a blood sample. Pulse oximetry estimates arterial oxygen saturation from optical signals, with Aoyagi's work contributing an important measurement principle. Capnography records carbon dioxide in expired gas over time.[114–116] Differences between these observations may be informative because oxygenation, ventilation, and perfusion are not equivalent quantities.

Continuous recording makes transient events and trajectories available for analysis. It also records disturbances from motion, poor contact, line damping, electrosurgery, perfusion changes, and processing. An apparently reasonable value can be misleading when the signal quality is poor. Waveforms, acquisition conditions, and the patient's observed state provide supporting information. A monitor display should not be treated as a self-validating account of physiology.

The history of monitoring also includes unequal measurement performance. Sjoding and colleagues' 2020 study found more frequent occult hypoxemia in patients recorded as Black than in those recorded as White when pulse-oximeter readings were compared with arterial measurements.[117] The study used racial categories rather than direct measurement of pigmentation, so the categories should not be mistaken for a complete optical mechanism. Its diagnostic significance is the need to examine performance across relevant populations and reconcile a reassuring display with other evidence when they disagree.

Alarms convert selected conditions into requests for attention. Too many non-actionable alarms can reduce attention to later warnings, while poorly chosen variables or thresholds can leave deterioration unrecognized. Safe monitoring depends on settings, escalation, staffing, maintenance, and time synchronization as well as sensor performance. Reviewing whether an alarm led to an appropriate response examines the monitoring process at a higher level than reviewing whether the alarm was generated.

27 Ambulatory and wearable monitoring

Holter recording made cardiac electrical activity observable during ordinary daily life.[118] Intermittent events could then be related to activity and symptoms rather than inferred from a short clinic recording. Event recorders, implanted loop recorders, ambulatory blood-pressure monitoring, sleep studies, and continuous glucose monitoring developed related forms of observation. Their sampling intervals, recorded quantities, and triggering rules determine which events become available for interpretation.

Wearables added further combinations of motion, optical pulse signals, temperature, electrical measurements, and patient-entered events.[119] Frequent recording can reveal useful patterns, but wear time, fit, motion, skin characteristics, firmware, and proprietary processing affect the record. Missing intervals may coincide with the very activity or illness being investigated. A derived estimate of sleep, rhythm, or activity should remain distinguishable from the underlying sensor observation.

Remote monitoring creates a continuing analytical workload. Results must be assigned to the correct person, summarized, reviewed, and connected to a defined response. A threshold suitable for one phase of illness may be inappropriate later. Summaries can remove the transient event that motivated monitoring, while excessive unfiltered data can prevent timely review. The organization of attention becomes part of the evidence system.

Within-person comparison is particularly useful when repeated measurements establish a sufficiently stable baseline. A change may become apparent before a population threshold is reached. However, the comparison requires awareness of device changes, missing intervals, treatment, and ordinary variation. Consent, access to data, cybersecurity, and responsibility for review also matter. The person being monitored needs to know when the service is observing the stream and when it is not.

28 Medical records and diagnostic interoperability

The medical record preserves diagnostic history as well as current results. Weed's problem-oriented record organized findings, assessments, and plans around explicit problems.[120] This made unresolved questions easier to follow across encounters. Electronic records later supported search, aggregation, and decision support, but also introduced copied text, fragmented displays, and documentation shaped by administrative requirements. A recorded statement may therefore describe a prior interpretation rather than a newly established fact.

Interoperability has several levels. DICOM provides structures for imaging information; LOINC supplies identifiers for laboratory and clinical observations; SNOMED CT represents clinical concepts; ICD supports classification and statistics. HL7 messaging and FHIR provide ways to exchange health information.[82, 121–123] These resources have different purposes and version histories. Moving information through a standard interface does not guarantee that the receiving system preserves its original meaning.

Units, specimen, body site, reference interval, method, negation, uncertainty, and provenance can change the meaning of a result. Two similarly named measurements may not be equivalent. A preliminary result, a correction, and a final report are also different states of the diagnostic record. Preserving those states helps explain the decisions made at the time instead of silently replacing their evidence with a later version.

The record should connect the question, request, acquisition, result, interpretation, communication, and subsequent action. An expected result that never arrives is different from a negative result. An acknowledged report is different from completed follow-up. These distinctions allow a diagnostic process to be examined for discontinuities. Interoperability becomes useful when the unresolved diagnostic questions and their evidence remain understandable as care passes between people and organizations.

29 Screening and population surveillance

Population investigation makes relationships among cases available as evidence. Snow's cholera work compared deaths, places, water supplies, and exposures before the causal organism had been established.[124] His wider comparisons of water suppliers are important alongside the familiar Broad Street episode. Removal of the pump handle alone does not account for the investigation's causal argument. Case definitions, denominators, and comparisons were necessary to distinguish a local cluster from a more general explanation.

Longitudinal studies such as Framingham connected measured characteristics with subsequent outcomes.[125] Risk estimation could then inform prevention before symptoms appeared. However, a risk association is different from a diagnosis of an existing condition, and an equation developed in one population may require reassessment elsewhere. Changes in treatment, exposure, referral, and ascertainment can alter the relationship between a recorded factor and its later outcome.

Screening examines people without recognized symptoms of the target condition. The target prevalence and consequences of error therefore differ from many diagnostic referrals. Wilson and Jungner's principles connected testing with disease importance, an identifiable early stage, acceptable methods, treatment, facilities, and a continuing program.[126] Earlier detection alone does not establish benefit. Lead-time bias, preferential detection of slower disease, overdiagnosis, false positives, and unequal follow-up can all change the apparent result.

Surveillance extends this process to continuing population observation. Clinical reports, laboratory notifications, registries, syndromic indicators, sequencing, and environmental measurements provide different evidence. Their coverage depends on who is tested, how a case is defined, and which results are reported. WHO's work on classification and diagnostic access reflects the need for comparable observations and practical testing capacity.[123, 127] A population signal should be connected to an appropriate public response without treating every individual as fully observed.

30 Diagnostic performance and safety

Formal evaluation of diagnostic tests separates sensitivity, specificity, predictive values, and likelihood ratios. Yerushalmy examined paired measures of diagnostic performance, while Ledley and Lusted related medical reasoning to probability. Fagan's nomogram provided a practical way to perform Bayesian updating.[128–130] These methods clarify why the same result can have different implications when the probability before testing is different.

Consider a hypothetical example with 1,000 people, of whom 10 have the target condition. A test with 90 percent sensitivity identifies 9 of those 10. At 90 percent specificity, it also gives positive results for 99 of the 990 people without the condition. Only 9 of the 108 positive results therefore come from people with the condition, approximately 8.3 percent. The example is an arithmetic illustration, not a performance claim about a clinical test. It shows why sensitivity is not the probability of disease after a positive result.

Performance also depends on which people and alternatives are included in the evaluation. Obvious disease compared with healthy controls may be easier to distinguish than the cases encountered in practice. If only selected participants receive the reference examination, the estimated performance can be biased. The reference standard itself can also be incomplete or wrong. The evaluation needs to preserve these conditions instead of reducing them to a single accuracy number.[131, 132]

Laboratory quality control examines the measurement process across repeated observations. Levey–Jennings charts and Westgard multirules made shifts, trends, and other patterns available for systematic investigation.[133, 134] External assessment adds comparison with other laboratories. Quality management and accreditation extend attention to competence and the wider examination process.[41] These activities detect problems that may be difficult to infer from any one patient's result.

Diagnostic safety concerns the process through which an explanation is established and communicated. The National Academies' 2015 report examined failures of accuracy, timeliness, and communication.[135] A technically correct test can be defeated by an incomplete history, a mistaken identity, an inaccessible service, a handoff failure, or missing follow-up. The patient and the people supporting them can also identify gaps that are absent from the formal record.

STARD and QUADAS-2 made important aspects of study reporting and bias assessment explicit.[131, 132] They help readers examine participant selection, the index test, the reference standard, and the sequence of verification. Reporting a method's best result without these conditions prevents meaningful comparison. The question is how well the evidence supports a stated use, not whether the test can be described as accurate in isolation.

A threshold connects a measurement to a decision. The boundary used for initial triage may differ from that used before an invasive procedure. Delay, missed disease, unnecessary intervention, available capacity, and patient preferences affect the consequences. A single statistically selected cutoff cannot

represent all of these purposes. Evaluation should specify both the intended action and the alternatives available after each result.

Uncertainty can be represented explicitly. An explanation may remain provisional because the evidence is incomplete, a finding has several plausible causes, or the condition is still evolving. The report can preserve the important alternatives and the circumstances that would require reassessment. Categorical wording should not conceal these limits. Follow-up then becomes part of the diagnostic process rather than evidence that the initial investigation somehow failed to reach finality.

Review of a diagnostic failure should reconstruct the information available at each stage. Later knowledge can make an earlier decision appear more certain than it was. Conversely, invoking uncertainty should not conceal a preventable failure to acquire, communicate, or act on evidence. Recurrent problems can be grouped and examined for mechanisms, followed by corrective changes and observation of their effects. This is diagnostic analysis applied to the diagnostic process itself.

PART VI

From evidence to computational diagnosis

31 Clinical decision support and expert systems

Computational decision support developed before contemporary machine learning. Probabilistic methods made the relationship between evidence and alternative explanations explicit, while decision analysis related those alternatives to possible actions.[129, 130] Computers could calculate scores, retrieve information, apply rules, and check selected constraints. The system needed a representation of the problem before it could contribute to its solution.

MYCIN, developed during the 1970s, represented knowledge about serious bacterial infections and antimicrobial selection through production rules and certainty factors.[136] It could show which rules contributed to a recommendation. Its historical importance should be separated from an assumption of routine clinical deployment. The project exposed persistent problems of knowledge acquisition, maintenance, validation, workflow, responsibility, and the meaning of an explanation produced by a machine.

An explicit rule can be inspected, but its scope may be narrow. Inputs have to represent the relevant findings, rules require maintenance, and exceptions may fall outside the knowledge base. Electronic-record alerts later placed related rule-based support directly into clinical work. An alert may prevent an error or add to an already excessive demand for attention. Its effect depends on timing, relevance, and the available response.

We should therefore evaluate decision support as an intervention in a process. A technically correct suggestion can arrive too late, be directed to the wrong person, or omit information needed for action. Responsibility for accepting, questioning, or overriding the suggestion has to be clear. The useful output is a supported decision in context, with a record that permits subsequent review.

32 Computer aided diagnosis

Computer-aided diagnosis in imaging and waveform analysis initially relied heavily on features selected by designers. Measurements of shape, intensity, texture, and timing were used to identify or classify possible abnormalities. Digital acquisition made these procedures more practical and reproducible.[137] The resulting process separated, at least conceptually, acquisition, feature construction, classification, and presentation to the reader.

Computer-aided detection often identifies a candidate region for further examination. It does not necessarily establish a diagnosis. A prompt may draw attention to a missed lesion, but it may also distract

the reader or generate unnecessary investigation. The effect depends on how the reader uses the output. Performance of the software on stored cases cannot establish the performance of a person using it during a different workflow.

Training and evaluation labels also require a source. They may come from pathology, reports, follow-up, a panel, or agreement among readers. Each has limitations and may describe a different target. A dataset can contain patterns associated with a scanner, hospital, referral route, or documentation habit instead of the intended condition. Those associations can disappear when the system is used elsewhere.

Evaluation therefore needs to examine the combined process. Relevant observations include missed disease, false-positive work, time, downstream procedures, disagreements, and patient outcomes. Performance across subgroups and settings may differ even when an overall metric looks satisfactory. Source evidence should remain available for inspection, and inconclusive or unsupported inputs should be distinguishable from reassuring findings.[137–139]

33 Machine learning and diagnostic AI

Deep learning changed how features could be obtained from diagnostic data. Studies of skin-lesion photographs and retinal images demonstrated strong performance on selected classification tasks.[138, 139] Such studies supported further work across imaging, pathology, signals, and molecular data. Comparison with a specialist benchmark remained conditional on the dataset, target, and evaluation procedure. It did not by itself demonstrate equivalence across the full work of a clinical specialty.

Prospective evaluation added another stage. Abràmoff and colleagues' 2018 trial examined an autonomous system for detecting diabetic retinopathy in primary-care offices against a defined reference examination.[140] This moved evaluation beyond classification of a convenient stored image set. Acquisition, image quality, operators, eligibility, and referral were part of the intended use. A system validated for one bounded task should not be interpreted as an autonomous general diagnostician.

Machine-learning results depend on data construction, labels, prevalence, devices, missingness, and the setting of care. A model can recognize useful correlations without representing the mechanism of disease. Some correlations may be incidental to how the data were acquired. Changes in population or practice can then alter discrimination or calibration. Validation needs to address the setting in which the output will influence decisions.

Multimodal models combine forms of evidence such as text, images, waveforms, and structured results. Generative systems can also summarize records or propose possible explanations. These capabilities introduce errors of their own, including invented details, loss of qualifiers, and unsupported causal connections. A fluent diagnostic narrative is not an independent source of evidence. Its claims must remain connected to the observations from which they were derived.

Regulatory and institutional guidance increasingly addressed diagnostic AI as a continuing activity. The FDA maintains a public resource on authorized AI-enabled medical devices, while WHO's 2021 guidance examines ethics and governance.[141, 142] Authorization and guidance have defined scopes; they should not be treated as evidence that every deployment is effective. Changes in software, population, interface, or intended use can alter the process after its initial evaluation.

An inspectable AI-assisted process preserves the intended use, input quality, software version, output, uncertainty, and subsequent human action. Overrides and disagreements are also useful evidence. They can reveal a limitation in the model, a problem in acquisition, or a gap in the surrounding workflow. Responsibility remains with the organizations and people providing the service, even when a bounded task is performed autonomously.

Monitoring after deployment should include calibration, subgroup performance, missing inputs, abstentions, downstream testing, and outcomes. A scanner upgrade, changed assay, new referral pattern, or revised documentation template can change the input distribution. Average performance may hide a new failure in a small but important group. Version control and change review connect an observed change to the system that produced it and support revision or withdrawal when necessary.

Combining a person with an AI system does not guarantee better performance. Readers may over-trust a suggestion or disregard correct output after repeated false alarms. They may also need time to resolve low-value disagreements. Evaluation should preserve the actual order of presentation and the information available to each participant. The outcome of interest is the effect on the diagnostic process and the patient, not an assumed benefit from having two interpreters.

Wider integration of records, images, laboratory results, genomes, and personal monitoring can reveal patterns across time and modalities. It also increases the effect of mistaken identity and shared data errors. Privacy-preserving methods may reduce some forms of data movement but do not automatically establish consent, representation, or fair access. More comprehensive data collection still requires a defined purpose and examination of its consequences.

Some computational applications are useful because they address a specific gap: detecting poor acquisition, locating missing results, retrieving an earlier study, prioritizing a review, or making uncertainty visible. These applications can be evaluated against the process they are intended to improve. The development of a general conversational interface should not obscure the continued need for such concrete diagnostic functions and their verification.

PART VII

Case studies and diagnostic principles

34 Diagnostic lessons from medical investigations

Childbed fever

Semmelweis's investigation of childbed fever compared mortality in the maternity divisions of the Vienna General Hospital. He related the excess deaths in one division to practices that included movement from autopsy work to obstetric care. The introduction of chlorinated-lime hand cleansing in 1847 was followed by a substantial reduction in mortality.[143] The investigation connected a population difference, an exposure hypothesis, an intervention, and subsequent observation before an adequate microbiological explanation was available.

We can distinguish the observed improvement from the later account of its mechanism. Comparisons required denominators and attention to how patients and staff moved through the institution. Sustaining the improvement required changes in practice and communication, not only recognition of the initial pattern. The case is therefore more useful than a simple story of one correct observer opposed by everyone else. It also concerns the women whose illnesses and deaths supplied the evidence and whose care was the purpose of the investigation.

The Legionella outbreak investigation

The 1976 outbreak associated with an American Legion convention required investigators to determine which illnesses belonged to the same event. Case finding, interviews, timing, pathology, and environmental comparisons narrowed the possibilities. Laboratory work subsequently isolated the organism associated with the outbreak and connected it with other respiratory disease. The epidemic description and the organism investigation supplied different but complementary evidence.[144, 145]

Early negative results did not exclude an agent that the available methods were not yet recovering. Preserved specimens and revised laboratory procedures made further investigation possible. The case also shows the importance of connecting observations acquired by different organizations. A case definition, a stored specimen, and an environmental history become more informative when their identities and relationships are preserved. No one of them contains the entire explanation.

Peptic ulcer disease

Marshall and Warren's work related curved bacteria in gastric tissue to gastritis and peptic-ulcer disease.[146] Findings from tissue examination, culture, clinical observation, treatment, and subsequent

studies changed the causal account of many ulcers. The 2005 Nobel Prize recognized their contribution.[147] A recurrent finding that could previously be dismissed or interpreted within an established explanation became the subject of a different mechanistic investigation.

The revised explanation still had limits. Not every ulcer is caused by Helicobacter pylori, and treatment failure may require investigation of resistance, reinfection, another cause, or an incorrect initial account. A mechanism explains a defined set of cases under particular conditions. Recognizing that scope is part of diagnostic reasoning. Treatment response can strengthen an explanation, but its interpretation also requires attention to other interventions and the natural course of the condition.

Coronavirus diagnostics

The availability of the novel coronavirus sequence in January 2020 supported rapid development of RT-PCR procedures.[148] Subsequent testing used laboratory and cartridge nucleic-acid assays, antigen detection, serology, sequencing, and environmental surveillance. These methods observed different targets. A result about viral RNA, a viral protein, an immune response, or a population wastewater signal could not be substituted directly for the others.

The time of collection and the purpose of testing affected interpretation. Sampling and stage of infection could influence detection; persistent RNA did not necessarily indicate ongoing transmissibility. Rapid antigen tests provided different sensitivity and timing tradeoffs, while serology addressed an immune response rather than serving as a general early-detection method. The pandemic also exposed failures of access, capacity, reporting, and follow-up. Producing a result was one part of the response, and its practical value depended on what support and action followed.

Comparison of the investigations

These investigations share an iterative structure. Observers acquired evidence, recognized a possible pattern, considered mechanisms, and sought observations that could distinguish them. An intervention or new method could change the available evidence and require another interpretation. Institutions determined whether the resulting knowledge was preserved and used. The diagnostic sequence therefore includes both the biological process being investigated and the history of the investigation itself.

35 Recurring principles and future directions

Start with the person and the clinical question

The clinical question determines what evidence is relevant. Screening, confirmation, localization, staging, prognosis, monitoring, and treatment selection have different requirements. The person's symptoms, priorities, circumstances, and preferences also affect the question. Acquiring a technically excellent result for a different purpose does not resolve the original problem.

Separate observation from interpretation

An observation and the conclusion drawn from it should remain distinguishable. The waveform, image, specimen description, or original account may support several interpretations. A later reviewer may need to examine that source evidence when a summary or label no longer fits. Preserving the relationship between source and interpretation makes revision possible.

Preserve identity provenance and time

The patient, specimen, site, method, time, units, and result status are supporting information rather than administrative decoration. Without them, otherwise correct measurements can be connected incorrectly. A corrected result should preserve its relationship to the earlier report so that subsequent analysis can reconstruct what was known when a decision was made.

Treat trajectories as evidence

Repeated observations reveal changes that a single encounter cannot show. Symptoms, measurements, and responses can be compared with earlier findings from the same person. The comparison must account for method changes, ordinary variation, treatment, and missing intervals. A trajectory is a sequence of observations with a context, not simply a line joining numbers.

Examine the reference standard and its limits

Pathology, culture, follow-up, and expert agreement can each provide a reference for a defined question. They differ in what they detect and in their possible errors. Evaluation should make the chosen reference and its application visible, including which participants did not receive it. Calling a method a gold standard does not remove these limitations.[131, 132]

Carry pretest probability forward

The probability before testing affects the interpretation afterward. Sensitivity and specificity describe conditional performance under an evaluation; they do not directly give the probability that a particular person has the condition after a result. The hypothetical example in Chapter 30 illustrates the difference. A test updates an account already informed by the history and setting.[128–130]

Distinguish measurement validity and clinical value

Reliable measurement, a useful relationship with disease, and improved care are separate achievements. A biomarker or algorithm may satisfy one without satisfying the others. Evaluation should identify which question has actually been answered and what remains to be established for the proposed use.

Combine independent evidence across modalities

Agreement is more informative when the contributing observations have different sources of error. History, physiology, imaging, pathology, and follow-up may constrain an explanation in complementary ways. Several reports copied from one interpretation do not constitute independent confirmation. The relationships among evidence sources should therefore be examined as well as their individual results.

Make missingness and measurement failure visible

An absent specimen, failed acquisition, uninterpretable result, and negative finding are different diagnostic states. They should not be represented by the same empty field or reassuring label. Making acquisition and communication failures visible allows unresolved questions to be followed rather than silently removed from the record.

Close the loop

The process continues after a result is generated. Interpretation, communication, agreed action, and follow-up determine whether the evidence contributes to care. A system that counts completed tests but cannot identify unresolved results has only observed part of the process. Closing the loop includes the patient's understanding and the practical completion of the next step.[135]

Learn from discordance

Discordance provides a reason to examine the current explanation. An autopsy, second opinion, later outcome, or patient report may reveal information missed earlier. Review should consider acquisition, interpretation, communication, and the conditions of care. Corrective action then needs its own follow-up so that its effect can be assessed.

Examine access and performance across settings

A test may be unavailable or perform differently in the setting where it is needed. Reference populations, language, disability access, cost, infrastructure, and subsequent treatment affect practical diagnostic value. Evaluating these differences is part of examining the intended use. A method's success at one institution cannot establish that every necessary condition is present elsewhere.

Make human interpretation open to review

Human participation must involve an intelligible opportunity to assess the evidence. The person reviewing a recommendation needs relevant source information, uncertainty, and alternatives. The organization also needs a clear account of who may question the result and who responds. A routine approval action alone does not demonstrate effective oversight.[141, 142]

Manage diagnostic data and AI throughout use

Diagnostic software and AI require continuing management of data, versions, performance, and reported problems. Changes in the population or workflow can invalidate earlier assumptions without a change in the model itself. Review should identify when revalidation, restriction, or retirement is needed. Privacy and consent remain relevant throughout the use and reuse of the data.

Preserve different forms of evidence

Different evidence types preserve different aspects of a person's condition. Combining them can improve an explanation without making it complete. The patient's account remains relevant when measurements are extensive and when measurements are normal. An adequate diagnostic process can preserve uncertainty while still supporting appropriate action and subsequent reassessment.

Artifacts Patterns and Mechanisms

In our pattern-oriented software diagnostics, we distinguish diagnostic artifacts, recognized patterns, and the mechanisms that may explain them.[149] A related analytical distinction helps us read this medical history. A blood film, ECG recording, image, or case narrative supplies evidence. A recurrent arrangement of findings supports pattern recognition. Physiology, pathology, microbiology, and other knowledge supply candidate mechanisms. Moving between these levels is iterative: an unexplained pattern can prompt a different acquisition, and a new result can require revision of the proposed mechanism.

Terminology must remain appropriate to the domain. In medicine, an artifact commonly means a feature introduced by acquisition or processing rather than by the target biological process. We therefore use diagnostic material or recorded evidence when referring generally to specimens, images, traces, and reports. The comparison with software concerns analytical relationships. It does not assert that the same pattern catalog or causal model can be applied unchanged to a patient.

For example, disagreement between a pulse-oximeter value and an arterial measurement can prompt examination of timing, acquisition, physiology, and device performance. At a second level, a review can ask why that disagreement was overlooked or why the result did not reach the responsible clinician. This is analysis of the diagnostic analysis. The history of the investigation can itself contain missing observations, premature conclusions, and unexamined assumptions.

Pattern names and checklists can help make such recurring activities explicit, but their value depends on what they preserve and what they leave out. A familiar pattern can arise through several mechanisms. An unfamiliar presentation can still involve a known mechanism. Medical use of a proposed analytical scheme would require appropriate clinical evaluation; the comparison made here is a way to organize historical understanding, not a claim of clinical validation.

Conclusion

We have followed changes in what medical practitioners could observe and how they could interpret it. Case histories preserved illness over time. Anatomy connected findings with organs and tissues. Laboratory methods examined specimens at cellular, microbial, chemical, and molecular levels. Imaging and physiological recording made selected internal structures and processes available during life. Population investigation and computation added further ways of comparing observations and constructing explanations.

Each development also introduced conditions that had to be understood. Instruments required suitable acquisition and calibration. Specimens required identification and handling. Images required knowledge of projection or reconstruction. Records required preservation of meaning and time. Models required an account of their inputs, assumptions, and scope. Diagnostic progress included making these conditions open to examination and correcting failures in the processes around them.

The historical examples also separate recognizing a pattern from establishing its mechanism. A recurrent sign can support several explanations, and an apparently inconsistent result may reveal a problem with either the working diagnosis or the method of observation. Additional evidence, comparison, intervention, and follow-up provide ways to examine those alternatives. The resulting explanation remains connected to the person whose condition is being investigated.

Medical diagnostics is therefore an iterative activity of acquisition, interpretation, explanation, and review. Its outcome includes communicating what is known, what remains uncertain, and what follows from that uncertainty. Observability supports this activity when it preserves the information needed to recognize change and reconsider an account. Debugging the diagnostic process contributes when it identifies and corrects the failures that prevent evidence from becoming effective care.

Appendix A Selected chronology

DATE	DEVELOPMENT	DIAGNOSTIC CONSEQUENCE
c. 1600 BCE	The surviving Edwin Smith Papyrus copies older trauma cases	Examination, judgment, and prognosis are organized as repeatable case forms.[1, 2]
c. 1100–700 BCE	Mesopotamian diagnostic and prognostic series are compiled	Signs are arranged into patterned judgments within a learned textual system.[3]
5th–4th c. BCE	Hippocratic case narratives circulate	Illness is followed through daily signs, regimen, crises, and outcome.[4]
c. 2nd c. BCE–2nd c. CE	Chinese medical classics take layered form	Pulse and other observations are integrated within a systematic medical cosmology.[5, 7]
2nd c. CE	Galenic writings synthesize anatomy, pulse, and humoral medicine	A highly transmissible diagnostic framework shapes Eurasian medicine for centuries.[5, 6, 14]
9th–10th c.	Al-Razi records and compares clinical cases	Bedside observation and therapeutic response are preserved in Islamic medicine.[8]
Early 11th c.	Ibn Sina organizes the Canon of Medicine	Medical principles, diagnosis, and treatment are organized for teaching and comparison.[9]
1543	Vesalius publishes De humani corporis fabrica	Printed, dissection-based anatomy challenges inherited authority.[10]
1628	Harvey publishes De Motu Cordis	Circulation is explained through anatomy, experiment, and quantitative reasoning.[11]
1660s–1670s	Microscopic investigations by Hooke, Malpighi, and van Leeuwenhoek expand	Cells, capillaries, and microorganisms become available to microscopic investigation.[28]
1761	Morgagni publishes De Sedibus	Symptoms during life are correlated systematically with postmortem organ lesions.[12]
1761	Auenbrugger publishes chest percussion	Sound elicited at the bedside becomes evidence of hidden thoracic density.[19]
1816–1819	Laennec develops and publishes mediate auscultation	Internal sounds become localizable and teachable through the stethoscope.[17]
1838–1839	Cell theory is articulated	Cells become a general unit for describing living structure.[29]
1846	Hutchinson publishes his spirometric investigation	Lung capacity becomes a measured functional observation related to bodily characteristics and disease.[26]
1847	Semmelweis introduces chlorinated-lime hand cleansing	Comparative mortality and intervention test a causal explanation of childbed fever.[143]
1851	Helmholtz introduces the ophthalmoscope	The living fundus becomes accessible to direct clinical observation.[80]
1854	John Snow investigates cholera around the Broad Street pump	Spatial and population comparison support a waterborne explanation.[124]
1858	Virchow publishes Cellular Pathology	Disease is reframed through changes at cellular scale.[13]
1860s–1870s	Pasteur and Koch advance germ theory and culture methods	Specific microorganisms become testable causal candidates.[30, 31]
1882	Koch announces the tubercle bacillus	Stain, microscopy, culture, and causal argument converge on a major disease.[31]
1884	Gram develops the differential bacterial stain	A rapid microscopic partition guides classification and initial treatment.[33]
1895	Röntgen discovers X-rays	Internal anatomy becomes visible without incision.[55]

DATE	DEVELOPMENT	DIAGNOSTIC CONSEQUENCE
1896–1905	Riva-Rocci's cuff and Korotkoff sounds make blood pressure practical	Arterial pressure becomes a routine quantified vital sign.[24, 25]
1900–1903	Landsteiner and colleagues describe the ABO groups; Einthoven refines the ECG	Transfusion compatibility and cardiac electrical diagnosis gain standard forms.[43, 58]
1906	The Wassermann reaction is introduced	Serological evidence enters diagnosis of syphilis.[150]
1924–1929	Berger records and publishes the human electroencephalogram	Brain electrical activity becomes a diagnostic trace.[59]
1941	Papanicolaou and Traut report diagnostic cervical cytology	Exfoliated cells become a basis for cancer detection and screening.[51, 52]
1945	The antiglobulin test is described	Incomplete red-cell antibodies become detectable.[44]
1952	Copenhagen reorganizes respiratory care during the polio epidemic	Concentrated monitoring and ventilation help shape intensive care.[113]
1956	Tjio and Levan establish the human chromosome number as 46	The correct count and improved preparations provide a firmer basis for chromosome diagnosis.[84]
1957	Skeggs describes continuous-flow analysis	Laboratory chemistry begins large-scale automation.[38]
1958	Abdominal and obstetric ultrasound investigations and the scintillation camera are reported	Non-ionizing imaging and planar radionuclide imaging expand.[62, 65]
1959	Trisomy 21 is linked to Down syndrome	A clinical syndrome receives a chromosomal mechanism.[85]
1961–1963	Guthrie screening and ambulatory ECG spread	Dried blood spots and Holter recording move diagnosis beyond specialist encounters.[86, 118]
1962	Clark and Lyons propose the enzyme electrode	Biosensing creates a path toward bedside and home glucose measurement.[109]
1960s	Radioimmunoassay reaches clinical use	Very low concentrations of hormones and other analytes become measurable.[46]
1971	ELISA is reported	Enzyme labels support scalable immunodiagnostics.[47]
1971–1973	Clinical CT and foundational MRI spatial encoding appear	Cross-sectional X-ray reconstruction and magnetic-resonance imaging emerge.[70–72, 75]
1975	Hybridoma methods produce monoclonal antibodies	Defined antibody reagents transform assays and tissue classification.[48]
1976–1977	Legionnaires' disease is investigated; MYCIN research matures	Outbreak systems identify a new pathogen while expert systems formalize rules.[136, 144, 145]
1977	Sanger chain-termination sequencing is published	DNA sequence becomes a scalable diagnostic evidence type.[95]
1978	Kemp reports stimulated acoustic emissions	Cochlear activity supplies a measurable acoustic response.[60]
1980s	PACS, ACR–NEMA imaging exchange, and digital modalities develop	Images become networked data with maintained identity and metadata.[82, 83]
1983–1985	PCR is conceived and demonstrated for human genetic targets	Selected nucleic acids can be amplified rapidly from scarce material.[91, 92]
1984	Marshall and Warren report gastric curved bacilli	Peptic-ulcer disease begins a major causal reframing.[146]

DATE	DEVELOPMENT	DIAGNOSTIC CONSEQUENCE
1990–1992	The Human Genome Project begins; BOLD contrast is described, then applied to human functional imaging	Genome-scale reference work and statistical functional imaging accelerate.[76, 77, 96–98]
1991	Huang and colleagues report optical coherence tomography	Optical interference permits cross-sectional tissue imaging at high resolution.[81]
1994	LOINC is created	Laboratory and clinical observations gain portable identifiers.[121]
1997	Cell-free fetal DNA is reported in maternal plasma	A route opens to non-invasive prenatal screening.[104]
1999	Vogelstein and Kinzler describe digital PCR	Partitioned reactions turn amplification into a counting-based measurement.[93]
2001	Draft human genome sequences are published	Genomic comparison becomes shared scientific infrastructure.[96, 97]
2003	Human Genome Project completion is announced	A major reference sequence and shared tools support sequencing; unresolved regions remain.[98, 101]
2005	Massively parallel pyrosequencing is reported	Next-generation sequencing expands throughput dramatically.[99]
2008	Counting cell-free DNA enables genome-wide aneuploidy screening	Prenatal screening shifts toward maternal plasma sequencing.[105]
2009	NIMH establishes the RDoC initiative	A research framework examines behavior and biology across existing diagnostic categories.[22]
2010s	Deep learning expands in medical imaging	Learned representations support computer-aided classification at scale.[138, 139]
2017	CRISPR-Cas13a nucleic-acid detection is reported	Programmable recognition and a reporter signal provide another route to molecular detection.[94]
2018	WHO introduces the Essential Diagnostics List	Diagnostic access and national test selection become explicit health-system priorities.[112]
2018	A pivotal autonomous retinal AI trial is published	Prospective evaluation examines a defined diagnostic task in primary-care offices.[140]
2020	Rapid SARS-CoV-2 RT-PCR protocols are published	Sequence sharing enables fast, distributed assay development during a pandemic.[148]
2021	WHO issues AI ethics and governance guidance for health	Diagnostic AI is framed as a lifecycle, rights, safety, and equity problem.[142]
2022	The T2T Consortium publishes the CHM13 genome sequence	Previously unresolved regions are assembled in a nearly complete human reference without a Y chromosome.[101]
2023	A draft human pangenome reference is published	Multiple diverse genome assemblies extend the representation used for genomic comparison.[102]
2024	WHO publishes ICD-11 clinical diagnostic requirements for mental disorders	Narrative, behavioral, and functional findings receive a standardized clinical description.[21]
2020s	Multimodal models and continuous sensing converge	Text, image, waveform, molecular, and longitudinal data can be analyzed together, intensifying governance needs.[141, 142]

Table 3. Selected developments. Dates distinguish documented milestones from wider adoption.

Appendix B Glossary

Accuracy. Agreement between a result and the target state under defined conditions; for classifiers, the overall proportion correct can be misleading when classes are imbalanced.

Analytical sensitivity. In diagnostic assay usage, the ability to detect a small amount of the target. A limit of detection concerns detection under stated conditions; a limit of quantification additionally concerns acceptable quantitative performance. This differs from clinical sensitivity.

Analytical specificity. The extent to which a method measures the intended analyte without interference or cross-reaction.

Analytical validity. Evidence that a test accurately and reliably measures the claimed analyte, sequence, signal, or feature.

Artifact. A feature produced by acquisition, preparation, processing, reconstruction, display, or handling rather than by the target biological state.

Assay. A defined procedure for detecting or quantifying an analyte or biological activity.

Auditory brainstem response (ABR). An electrical response recorded after sound stimulation to assess activity along the auditory pathway. Its interpretation depends on acquisition conditions and the clinical question.[61]

Bayesian updating. Revision of a prior probability using the evidential force of a new observation.

Bias. Systematic deviation introduced by design, sampling, measurement, analysis, interpretation, workflow, or unequal representation.

Biomarker. A defined characteristic measured as an indicator of biological processes, pathogenic processes, or responses to exposure or intervention.

Biosensor. A device coupling a biological recognition or response element to a transducer and readout.

Calibration. Establishing the relationship between instrument response and reference values under specified conditions.

Case definition. Operational criteria used to classify persons for surveillance or investigation; not necessarily identical to an individual clinical diagnosis.

Clinical decision support. Knowledge, rules, models, or information displays designed to assist a health decision within workflow.

Clinical utility. Evidence that use of a test or model improves decisions, outcomes, or other patient-relevant consequences.

Clinical validity. Evidence that a result is associated with or predicts the clinical condition or outcome of interest.

Cohort. A defined group followed or reconstructed across time to compare exposures, characteristics, or outcomes.

Computer-aided detection. Software that identifies candidate abnormalities for review, often without assigning a complete diagnosis.

Computer-aided diagnosis. Software that estimates, classifies, or otherwise assists interpretation of a diagnostic target.

Confounding. Distortion of an association by another factor related to both the observed exposure or feature and the outcome.

CRISPR diagnostic assay. A method using sequence recognition by a CRISPR-associated system to generate a detectable signal from a selected nucleic-acid target. Amplification, preparation, and readout may remain necessary.[94]

Cutoff or threshold. A boundary used to convert a continuous or graded result into categories or actions.

Diagnosis. An explanatory account of a health condition reached by integrating evidence, alternatives, probability, and context and communicating that account.

Diagnostic error. Failure to establish an accurate and timely explanation of a patient's health problem, or failure to communicate that explanation to the patient.[135] Diagnostic knowledge may subsequently change as new evidence becomes available.

Diagnostic pattern. A recurrent arrangement of findings that supports recognition and further analysis. The same pattern may have several mechanisms, and recognition alone does not establish causation.

DICOM. Digital Imaging and Communications in Medicine, the standard family for medical imaging objects, metadata, and network services.

Differential diagnosis. A structured set of plausible alternative explanations requiring comparison or exclusion.

Digital pathology. Acquisition, management, viewing, exchange, and computation using digitized pathology material.

Digital PCR. PCR performed across many separated reactions so that positive and negative partitions support estimation of target abundance under an explicit counting model.[93]

Drift. Change over time in an instrument, assay, population, data distribution, workflow, or model performance.

Electronic health record (EHR). A digital longitudinal record and workflow environment for health information; its content is partial and purpose-shaped.

ELISA. Enzyme-linked immunosorbent assay, using enzyme-mediated signal to detect antigen–antibody binding.

False negative. A negative result in a person who has the target condition under the chosen reference definition.

False positive. A positive result in a person who does not have the target condition under the chosen reference definition.

FHIR. Fast Healthcare Interoperability Resources, an HL7 standard for exchanging healthcare information electronically.[122]

FISH. Fluorescence in-situ hybridization, using labeled nucleic-acid probes to locate selected sequences in cells, chromosomes, or tissue.

Functional test. An examination of physiological performance or response under defined conditions. The task or stimulus, baseline, timing, and acquisition quality are part of the evidence.

Genotype. Genetic constitution at selected loci or genome scale; it does not determine phenotype independently of context.

Gold standard. Informal term for a best available reference method; potentially misleading because reference standards also have error and scope limits.

Hounsfield unit. A CT attenuation scale referenced to water and air under defined reconstruction and calibration conditions.

ICD. International Classification of Diseases, maintained by WHO for comparable health and mortality classification.[123]

Immunohistochemistry. Antibody-based detection of selected antigens while preserving tissue location and morphology.

Incidence. The occurrence of new events or cases in a population over a defined interval.

Incidental finding. An observation outside the original diagnostic purpose that may create new uncertainty or follow-up obligations.

Inference. A conclusion drawn from evidence and a model rather than directly observed as such.

Interoperability. The capacity of systems to exchange information and preserve usable meaning, context, identity, and status.

Karyotype. An organized representation of chromosome number and structure from selected cells.

Lead-time bias. An apparent increase in survival measured from diagnosis that results from detecting disease earlier, even when the time of death is unchanged.

Length bias. Preferential detection of slowly progressing disease by screening, making detected cases appear to have a better prognosis even without a treatment benefit.

Likelihood ratio. The probability of a test result among people with the target condition divided by its probability among people without it.

Liquid biopsy. Analysis of circulating cells, nucleic acids, vesicles, or other disease-derived material in a body fluid.

LOINC. A terminology providing identifiers for laboratory and clinical observations.[121]

Matrix effect. Influence of other specimen components on measurement of the intended analyte.

Measurand. The quantity intended to be measured, defined sufficiently to make results interpretable.

Mechanism. A process proposed to explain how an observed condition or pattern arises. Evidence of association or a familiar pattern does not by itself establish the mechanism.

Modality. A class of acquisition or examination, such as CT, MRI, ultrasound, pathology, or laboratory testing.

Monitoring. Repeated observation of selected variables to recognize change, risk, response, or need for action.

Negative predictive value. Probability of not having the target condition given a negative result; dependent on prevalence and selection.

Next-generation sequencing (NGS). Massively parallel sequencing methods capable of reading many DNA or RNA fragments in one run.

Observability. The capacity of an evidence system to support inferences about otherwise hidden physiological and pathological states. We use the term here as a retrospective analytical perspective; it does not imply complete visibility or deliberate historical design.

Optical coherence tomography (OCT). An imaging method that uses optical interference and differences in backscattered light to resolve tissue structure in depth.[81]

Otoacoustic emission (OAE). Sound arising from cochlear activity and measurable in the ear canal. An emission test assesses a different part of hearing function from a behavioral response or an auditory brainstem recording.[60, 61]

Overdiagnosis. Correct detection of a condition that would not have caused symptoms or harm during the person's lifetime, leading potentially to unnecessary intervention.

Pangenome reference. A representation of genomic sequences and variation from multiple individuals that supports comparison beyond a single linear reference sequence.[102]

Pathology. Study and diagnosis of disease through organs, tissues, cells, fluids, molecules, and clinicopathological correlation.

PCR. Polymerase chain reaction, a method for amplifying selected nucleic-acid targets.

Phenotype. Observable characteristics resulting from genetic, developmental, environmental, and temporal influences.

Point-of-care test. A test performed near the patient or decision rather than in a distant central laboratory.

Positive predictive value. Probability of having the target condition given a positive result; dependent on prevalence and selection.

Precision. Closeness of agreement among repeated measurements under defined conditions; distinct from accuracy.

Pretest probability. Estimated probability of the target condition before the result under consideration.

Prognosis. A reasoned estimate of future course, outcome, or response; distinct from identifying the present condition.

Provenance. Information about origin, identity, collection, transformation, version, and custody of evidence.

Quality control. Procedures using controls, rules, comparison, and review to detect unacceptable measurement performance.

Quality management system. The coordinated organization of personnel, procedures, equipment, documentation, assessment, and improvement needed to support reliable laboratory work.[41]

Reference interval. The distributional interval expected for a defined reference population and method; not a universal boundary between health and disease.

Reference standard. The method or composite used to determine the target condition when evaluating a diagnostic test.

Reliability. Consistency of findings or classifications under specified repeated observations. Agreement can be high even when an interpretation is systematically wrong.

Repeatability. Agreement among measurements repeated under closely specified conditions, such as the same method, operator, instrument, and short time interval.

Screening. Testing or assessment of people without recognized symptoms of the target condition to identify who may need further evaluation.

Sensitivity. Proportion of people with the target condition who receive a positive result at the chosen threshold.

Serial testing. Interpretation of repeated tests over time or in sequence, including strategies that require concordance or escalation.

Signal. A measured variation carrying information about a physical or biological process, mixed with noise and acquisition effects.

Specificity. Proportion of people without the target condition who receive a negative result at the chosen threshold.

Specimen. Material collected from a person for examination; its identity, site, timing, handling, and adequacy are part of the evidence.

Spectrum bias. Change in observed test performance caused by differences in disease severity, alternatives, or patient characteristics across study and use populations.

Staging. Assessment of disease extent, distribution, or severity within a defined classification system.

Surveillance. Systematic population-level collection, analysis, interpretation, and communication for awareness and action.

Titer. A semi-quantitative expression of concentration based on the greatest dilution meeting a defined reaction criterion.

Triage. Prioritization of attention or care based on urgency, risk, and available resources; not a complete diagnosis.

Uncertainty. The range of plausible values, classifications, explanations, or outcomes remaining after available evidence.

Validation. Demonstration that a method, system, or model performs adequately for an intended use and population.

Verification bias. Distortion created when receipt of the reference standard depends on the index-test result or other selected features.

Wearable. A body-worn sensor or device collecting physiological, behavioral, or contextual data intermittently or continuously.

Appendix C Evidence taxonomy

EVIDENCE CLASS	TYPICAL EXAMPLES	CHARACTERISTIC STRENGTH	CHARACTERISTIC LIMITATION
Patient narrative and context	Symptoms, chronology, exposures, function, goals	Whole-person meaning, time course, low latency	Recall, language, power, access, incomplete documentation
Physical signs	Appearance, palpation, percussion, sounds, function	Immediate, repeatable, integrates multiple systems	Observer variation, subtlety, environment, limited localization
Mental and cognitive assessment	Reported experience, speech, affect, memory, structured interview	Organizes experience, behavior, function, and time course	Language, setting, observer agreement, category heterogeneity
Vital and physiological measurements	Temperature, pressure, ECG, EEG, oximetry, capnography	Quantified, serial, alarmable	Artifact, calibration, thresholding, proxy relationships
Functional responses	Spirometry, task performance, physiological challenge	Shows performance under recorded test conditions	Effort, stimulus, baseline, timing, reference assumptions
Specimen chemistry and hematology	Electrolytes, enzymes, hormones, cell counts	Sensitive quantification, automation, trend comparison	Pre-analytics, interference, reference intervals, nonspecificity
Microbiological evidence	Stain, culture, antigen, NAAT, susceptibility	Organism-level causal and treatment information	Sampling, contamination, viability, colonization, time
Morphology and pathology	Cytology, biopsy, histology, autopsy	Architecture, cellular detail, clinicopathological revision	Sampling, preparation, observer variation, invasiveness
Anatomical imaging	Radiography, CT, MRI, ultrasound	Localization and extent inside the living body	Resolution, reconstruction, incidental findings, operator/protocol effects
Functional and molecular imaging	PET, SPECT, perfusion, fMRI	Physiology or metabolism with spatial context	Indirectness, limited specificity, modeling and timing assumptions
Genetic and genomic evidence	Karyotype, FISH, panel, exome, genome	Inherited or acquired molecular mechanism	Uncertain variants, ancestry bias, reclassification, consent implications
Longitudinal and wearable evidence	Holter, continuous glucose, remote sensors, symptom diaries	Trajectories and events outside encounters	Missingness, proprietary processing, response burden, privacy
Population evidence	Cohorts, registries, screening, outbreak surveillance	Denominators, comparison, general patterns	Selection, confounding, ecological inference, transportability
Therapeutic response	Improvement after targeted intervention	Can test mechanism within a time course	Placebo, regression, co-intervention, nonspecific response
Computational inference	Risk score, rules, CAD, machine learning, multimodal AI	Scalable integration and pattern recognition	Dataset shift, opacity, shortcut learning, automation bias
Diagnostic review and follow-up	Second opinion, discordance review, autopsy, outcomes	Additional evidence for correction and learning	Delay, incomplete ascertainment, institutional and legal constraints
Evidence about diagnostic processes	Control charts, proficiency testing, result tracking, version records	Makes acquisition and workflow failures available for review	Incomplete records, unsuitable metrics, lack of follow-up

Table 4. Evidence classes, their uses, and their limitations. Agreement is more informative when sources of error are independent.

References

[1] Sanchez GM, Burridge AL. Decision making in head injury management in the Edwin Smith Papyrus. *Neurosurgical Focus*. 2007;23(1):E5. doi:10.3171/FOC-07/07/E5. PubMed

[2] van Middendorp JJ, Sanchez GM, Burridge AL. The Edwin Smith Papyrus: a clinical reappraisal of the oldest known document on spinal injuries. *European Spine Journal*. 2010;19:1815–1823. Full text

[3] Retief FP, Cilliers L. Mesopotamian medicine. *South African Medical Journal*. 2007;97(1):27–30. PubMed

[4] Hippocrates. *Of the Epidemics*, Books I and III. Translated by Francis Adams. Accessed 1 August 2026. MIT Classics Archive

[5] Hajar R. The pulse in ancient medicine, part 1. *Heart Views*. 2018;19(1):36–43. Full text

[6] Hajar R. The pulse in medieval and Arab-Islamic medicine: Part 2. *Heart Views*. 2018;19(2):76–80. doi:10.4103/heartviews.heartviews_100_18. Full text

[7] Wang YYL, Wang SH, Jan MY, Wang WK. Past, present, and future of the pulse examination. *Journal of Traditional and Complementary Medicine*. 2012;2(3):164–185. Full text

[8] Tibi S. Al-Razi and Islamic medicine in the 9th century. *Journal of the Royal Society of Medicine*. 2006;99(4):206–207. doi:10.1177/014107680609900425. Full text

[9] Nasser M, Tibi A, Savage-Smith E. Ibn Sina's Canon of Medicine: 11th century rules for assessing the effects of drugs. *Journal of the Royal Society of Medicine*. 2009;102(2):78–80. doi:10.1258/jrsm.2008.08k040. DOI; Earlier full-text commentary

[10] Mesquita ET, de Souza Júnior CV, Ferreira TR. Andreas Vesalius 500 years: a Renaissance that revolutionized cardiovascular knowledge. *Revista Brasileira de Cirurgia Cardiovascular*. 2015;30(2):260–265. Full text

[11] Harvey W. *An Anatomical Disquisition on the Motion of the Heart and Blood in Animals* (1628). Fordham Internet History Sourcebooks

[12] Ghosh SK. Giovanni Battista Morgagni (1682–1771): father of pathologic anatomy and pioneer of modern medicine. *Anatomical Science International*. 2017;92(3):305–312. PubMed

[13] Virchow R. *Cellular Pathology as Based upon Physiological and Pathological Histology*. 1858/1860. Internet Archive

[14] van den Tweel JG, Taylor CR. A brief history of pathology: preface to a forthcoming series that highlights milestones in the evolution of pathology as a discipline. *Virchows Archiv*. 2010;457(1):3–10. doi:10.1007/s00428-010-0934-4. Full text

[15] Walker HK. The origins of the history and physical examination. In: *Clinical Methods*. 3rd ed. Butterworths; 1990. NCBI Bookshelf

[16] Roguin A. Rene Theophile Hyacinthe Laënnec (1781–1826): the man behind the stethoscope. *Clinical Medicine & Research*. 2006;4(3):230–235. Full text

[17] Sakula A. R. T. H. Laënnec 1781–1826, his life and work: a bicentenary appreciation. *Thorax*. 1981;36(2):81–90. Full text

[18] Verghese A, Charlton B, Kassirer JP, Ramsey M, Ioannidis JPA. Inadequacies of physical examination as a cause of medical errors and adverse events: a collection of vignettes. *The American Journal of Medicine*. 2015;128(12):1322–1324.e3. doi:10.1016/j.amjmed.2015.06.004. PubMed

[19] Auenbrugger L. *Inventum novum ex percussione thoracis humani ut signo abstrusos interni pectoris morbos detegendi*. Vienna: Trattner; 1761. Internet Archive

[20] Reiser SJ. *Medicine and the Reign of Technology*. Cambridge University Press; 1978.

[21] World Health Organization. *Clinical descriptions and diagnostic requirements for ICD-11 mental, behavioural and neurodevelopmental disorders*. Geneva: WHO; 2024. ISBN 978-92-4-007726-3. WHO

[22] National Institute of Mental Health. About RDoC. Accessed 18 September 2026. NIMH

[23] Mackowiak PA, Wasserman SS, Levine MM. A critical appraisal of 98.6°F, the upper limit of the normal body temperature, and other legacies of Carl Reinhold August Wunderlich. *JAMA*. 1992;268:1578–1580. doi:10.1001/jama.1992.03490120092034. DOI

[24] Kotchen TA. Historical trends and milestones in hypertension research: a model of the process of translational research. *Hypertension*. 2011;58(4):522–538. doi:10.1161/HYPERTENSIONAHA.111.177766. DOI

[25] Beevers G, Lip GYH, O'Brien E. ABC of hypertension: Blood pressure measurement. Part II—Conventional sphygmomanometry: technique of auscultatory blood pressure measurement. *BMJ*. 2001;322(7293):1043–1047. doi:10.1136/bmj.322.7293.1043. Full text

[26] Hutchinson J. On the capacity of the lungs, and on the respiratory functions, with a view of establishing a precise and easy method of detecting disease by the spirometer. *Medico-Chirurgical Transactions*. 1846;29:137–252. doi:10.1177/095952874602900113. Full text

[27] Graham BL, Steenbruggen I, Miller MR, et al. Standardization of Spirometry 2019 Update. An Official American Thoracic Society and European Respiratory Society Technical Statement. *American Journal of Respiratory and Critical Care Medicine*. 2019;200(8):e70–e88. doi:10.1164/rccm.201908-1590ST. Full text

[28] Wollman AJM, Nudd R, Hedlund EG, Leake MC. From Animaculum to single molecules: 300 years of the light microscope. *Open Biology*. 2015;5(4):150019. doi:10.1098/rsob.150019. Full text

[29] Mazzarello P. A unifying concept: the history of cell theory. *Nature Cell Biology*. 1999;1(1):E13–E15. doi:10.1038/8964. DOI

[30] National Research Council (US) Committee to Update *Science, Medicine, and Animals*. A theory of germs. In: *Science, Medicine, and Animals*. Washington, DC: National Academies Press; 2004. NCBI Bookshelf

[31] Gradmann C. Robert Koch and the pressures of scientific research: tuberculosis and tuberculin. *Medical History*. 2001;45(1):1–32. Full text

[32] Evans AS. Causation and disease: the Henle–Koch postulates revisited. *Yale Journal of Biology and Medicine*. 1976;49(2):175–195. PMID 782050. Full text

[33] Bartholomew JW, Mittwer T. The Gram stain. *Bacteriological Reviews*. 1952;16:1–29. Full text

[34] Centers for Disease Control and Prevention. Microbiology Series. *CDC Laboratory Training*. Updated 23 September 2024. Accessed 1 August 2026. CDC

[35] Miller JM, Binnicker MJ, Campbell S, et al. A guide to utilization of the microbiology laboratory for diagnosis of infectious diseases: 2018 update by the Infectious Diseases Society of America and the American Society for Microbiology. *Clinical Infectious Diseases*. 2018;67(6):e1–e94. doi:10.1093/cid/ciy381. Full text

[36] Jorgensen JH, Ferraro MJ. Antimicrobial susceptibility testing: a review of general principles and contemporary practices. *Clinical Infectious Diseases*. 2009;49(11):1749–1755. doi:10.1086/647952. DOI; PubMed

[37] Kricka LJ, Savory J. International Year of Chemistry 2011: a guide to the history of clinical chemistry. *Clinical Chemistry*. 2011;57(8):1118–1126. doi:10.1373/clinchem.2011.165233. PubMed

[38] Skeggs LT Jr. An automatic method for colorimetric analysis. *American Journal of Clinical Pathology*. 1957;28(3):311–322. doi:10.1093/ajcp/28.3_ts.311. DOI

[39] Peterson HE, Jungner I. The history of the AutoChemist: from vision to reality. *Yearbook of Medical Informatics*. 2014;9(1):235–243. Full text

[40] Armbruster DA, Overcash DR, Reyes J. Clinical chemistry laboratory automation in the 21st century—Amat Victoria curam (victory loves careful preparation). *Clinical Biochemist Reviews*. 2014;35(3):143–153. PMID 25336760. Full text

[41] World Health Organization. *Laboratory quality management system: handbook*. Geneva: WHO; 2011. ISBN 978-92-4-154827-4. WHO

[42] Bain BJ. *Blood Cells: A Practical Guide*. 5th ed. Wiley-Blackwell; 2015. ISBN 978-1-118-81733-9 (print); 978-1-118-81729-2 (e-book). Publisher record

[43] Nobel Prize Outreach. Karl Landsteiner – facts. Nobel Prize in Physiology or Medicine 1930. Accessed 19 September 2026. NobelPrize.org

[44] Coombs RRA, Mourant AE, Race RR. A new test for the detection of weak and 'incomplete' Rh agglutinins. *British Journal of Experimental Pathology*. 1945;26(4):255–266. Full text

[45] Papp JR, Park IU, Fakile Y, Pereira L, Pillay A, Bolan GA. CDC laboratory recommendations for syphilis testing, United States, 2024. *MMWR Recommendations and Reports*. 2024;73(1):1–32. doi:10.15585/mmwr.rr7301a1. CDC

[46] Yalow RS, Berson SA. Immunoassay of endogenous plasma insulin in man. *Journal of Clinical Investigation*. 1960;39(7):1157–1175. doi:10.1172/JCI104130. DOI

[47] Engvall E, Perlmann P. Enzyme-linked immunosorbent assay (ELISA): quantitative assay of immunoglobulin G. *Immunochemistry*. 1971;8:871–874. PubMed

[48] Köhler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*. 1975;256:495–497. PubMed

[49] Coons AH, Creech HJ, Jones RN. Immunological properties of an antibody containing a fluorescent group. *Proceedings of the Society for Experimental Biology and Medicine*. 1941;47:200–202. doi:10.3181/00379727-47-13084P. DOI

[50] Jaafar H. Intra-operative frozen section consultation: concepts, applications and limitations. *Malaysian Journal of Medical Sciences*. 2006;13(1):4–12. PMID 22589584. Full text

[51] Papanicolaou GN, Traut HF. The diagnostic value of vaginal smears in carcinoma of the uterus. *American Journal of Obstetrics and Gynecology*. 1941;42(2):193–206. doi:10.1016/S0002-9378(16)40621-6. DOI

[52] National Cancer Institute. Cervical cancer screening. National Cancer Institute. Updated 13 February 2025. Accessed 1 August 2026. NCI

[53] Evans AJ, Brown RW, Bui MM, et al. Validating whole slide imaging systems for diagnostic purposes in pathology: guideline update from the College of American Pathologists in collaboration with the American Society for Clinical Pathology and the Association for Pathology Informatics. *Archives of Pathology & Laboratory Medicine*. 2022;146(4):440–450. doi:10.5858/arpa.2020-0723-CP. CAP

[54] Hanna MG, Ardon O. Digital pathology systems enabling quality patient care. *Genes, Chromosomes & Cancer*. 2023;62(11):685–697. doi:10.1002/gcc.23192. Full text

[55] Nobel Prize Outreach. Wilhelm Conrad Röntgen – facts. Nobel Prize in Physics 1901. Accessed 19 September 2026. NobelPrize.org

[56] Glasser O. *Wilhelm Conrad Röntgen and the Early History of the Roentgen Rays*. Springfield, IL: Charles C Thomas; 1934. Internet Archive record

[57] International Commission on Radiological Protection. Radiological protection in medicine. ICRP Publication 105. *Annals of the ICRP*. 2007;37(6). ICRP

[58] Nobel Prize Outreach. Willem Einthoven – facts. Nobel Prize in Physiology or Medicine 1924. Accessed 19 September 2026. NobelPrize.org

[59] Collura TF. History and evolution of electroencephalographic instruments and techniques. *Journal of Clinical Neurophysiology*. 1993;10:476–504. PubMed

[60] Kemp DT. Stimulated acoustic emissions from within the human auditory system. *Journal of the Acoustical Society of America*. 1978;64(5):1386–1391. doi:10.1121/1.382104. DOI

[61] National Institute on Deafness and Other Communication Disorders. Your baby's hearing screening and next steps. NIH Publication No. 21–4968. October 2021; updated 10 June 2024. Accessed 18 September 2026. NIDCD

[62] Donald I, MacVicar J, Brown TG. Investigation of abdominal masses by pulsed ultrasound. *Lancet*. 1958;271:1188–1195. doi:10.1016/S0140-6736(58)91905-6. DOI

[63] Singh S, Goyal A. The origin of echocardiography: a tribute to Inge Edler. *Texas Heart Institute Journal*. 2007;34(4):431–438. Full text

[64] Nobel Prize Outreach. George de Hevesy – facts. Nobel Prize in Chemistry 1943. Accessed 19 September 2026. NobelPrize.org

[65] Anger HO. Scintillation camera. *Review of Scientific Instruments*. 1958;29:27–33. doi:10.1063/1.1715998. DOI

[66] Kuhl DE, Edwards RQ. Image separation radioisotope scanning. *Radiology*. 1963;80(4):653–662. doi:10.1148/80.4.653. DOI

[67] Hutton BF. The contribution of medical physics to nuclear medicine: looking back—a physicist's perspective. *EJNMMI Physics*. 2014;1(1):2. doi:10.1186/2197-7364-1-2. Full text

[68] Jones T, Townsend D. History and future technical innovation in positron emission tomography. *Journal of Medical Imaging*. 2017;4(1):011013. doi:10.1117/1.JMI.4.1.011013. Full text

[69] Cormack AM. Representation of a function by its line integrals, with some radiological applications. *Journal of Applied Physics*. 1963;34:2722–2727. doi:10.1063/1.1729798. DOI

[70] Hounsfield GN. Computerized transverse axial scanning (tomography): part 1. Description of system. *British Journal of Radiology*. 1973;46:1016–1022. doi:10.1259/0007-1285-46-552-1016. DOI

[71] Nobel Prize Outreach. The Nobel Prize in Physiology or Medicine 1979: Cormack and Hounsfield. Accessed 19 September 2026. NobelPrize.org

[72] Lauterbur PC. Image formation by induced local interactions: examples employing nuclear magnetic resonance. *Nature*. 1973;242:190–191. doi:10.1038/242190a0. DOI

[73] Mansfield P, Maudsley AA. Medical imaging by NMR. *British Journal of Radiology*. 1977;50:188–194. doi:10.1259/0007-1285-50-591-188. DOI

[74] Mansfield P. Multi-planar image formation using NMR spin echoes. *Journal of Physics C: Solid State Physics*. 1977;10(3):L55–L58. doi:10.1088/0022-3719/10/3/004. DOI

[75] Nobel Prize Outreach. The Nobel Prize in Physiology or Medicine 2003: Lauterbur and Mansfield. Accessed 19 September 2026. NobelPrize.org

[76] Ogawa S, Lee TM, Kay AR, Tank DW. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*. 1990;87:9868–9872. Full text

[77] Kwong KK, Belliveau JW, Chesler DA, et al. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences of the USA*. 1992;89(12):5675–5679. doi:10.1073/pnas.89.12.5675. PubMed; Full text

[78] Spaner SJ, Warnock GL. A brief history of endoscopy, laparoscopy, and laparoscopic surgery. *Journal of Laparoendoscopic & Advanced Surgical Techniques*. 1997;7(6):369–373. PubMed

[79] Hirschowitz BI, Curtiss LE, Peters CW, Pollard HM. Demonstration of a new gastroscope, the "fiberscope." *Gastroenterology*. 1958;35:50–53. PubMed

[80] Keeler CR. The ophthalmoscope in the lifetime of Hermann von Helmholtz. *Archives of Ophthalmology*. 2002;120(2):194–201. doi:10.1001/archopht.120.2.194. Publisher

[81] Huang D, Swanson EA, Lin CP, et al. Optical coherence tomography. *Science*. 1991;254(5035):1178–1181. doi:10.1126/science.1957169. Full text

[82] DICOM Standards Committee. The history of DICOM. Accessed 1 August 2026. Official history

[83] Huang HK. Short history of PACS. Part I: USA. *European Journal of Radiology*. 2011;78(2):163–176. doi:10.1016/j.ejrad.2010.05.007. PubMed

[84] Tjio JH, Levan A. The chromosome number of man. *Hereditas*. 1956;42:1–6. doi:10.1111/j.1601-5223.1956.tb03010.x. DOI

[85] Lejeune J, Gautier M, Turpin R. Étude des chromosomes somatiques de neuf enfants mongoliens. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*. 1959;248(11):1721–1722. PubMed record

[86] la Marca G, et al. Current state and innovations in newborn screening: continuing to do good and avoid harm. *International Journal of Neonatal Screening*. 2023;9(1):15. PubMed

[87] Centers for Disease Control and Prevention. CDC Grand Rounds: newborn screening and improved outcomes. *MMWR Morbidity and Mortality Weekly Report*. 2012;61(21):390–393. CDC

[88] Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*. 1975;98:503–517. PubMed

[89] Bauman JGJ, Wiegant J, Borst P, van Duijn P. A new method for fluorescence microscopical localization of specific DNA sequences by in situ hybridization of fluorochrome-labelled RNA. *Experimental Cell Research*. 1980;128(2):485–490. doi:10.1016/0014-4827(80)90087-7. DOI; PubMed

[90] Mullis KB, Faloona FA. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology*. 1987;155:335–350. PubMed

[91] Mullis KB. The unusual origin of the polymerase chain reaction. *Scientific American*. 1990;262(4):56–61, 64–65. doi:10.1038/scientificamerican0490-56. DOI; PubMed

[92] Saiki RK, et al. Enzymatic amplification of beta-globin genomic sequences and restriction-site analysis for diagnosis of sickle cell anemia. *Science*. 1985;230:1350–1354. PubMed

[93] Vogelstein B, Kinzler KW. Digital PCR. *Proceedings of the National Academy of Sciences of the USA*. 1999;96(16):9236–9241. doi:10.1073/pnas.96.16.9236. Publisher

[94] Gootenberg JS, Abudayyeh OO, Lee JW, et al. Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science*. 2017;356(6336):438–442. doi:10.1126/science.aam9321. Full text

[95] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*. 1977;74:5463–5467. PubMed

[96] International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*. 2001;409:860–921. doi:10.1038/35057062. DOI

[97] Venter JC, et al. The sequence of the human genome. *Science*. 2001;291:1304–1351. doi:10.1126/science.1058040. DOI

[98] National Human Genome Research Institute. The Human Genome Project. Accessed 1 August 2026. NHGRI

[99] Margulies M, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2005;437:376–380. PubMed

[100] Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*. 2015;17(5):405–424. doi:10.1038/gim.2015.30. Full text

[101] Nurk S, Koren S, Rhie A, et al. The complete sequence of a human genome. *Science*. 2022;376(6588):44–53. doi:10.1126/science.abj6987. Full text

[102] Liao WW, Asri M, Ebler J, et al. A draft human pangenome reference. *Nature*. 2023;617(7960):312–324. doi:10.1038/s41586-023-05896-x. DOI; Author publication record

[103] Carlson LM, Vora NL. Prenatal diagnosis: screening and diagnostic tools. *Obstetrics and Gynecology Clinics of North America*. 2017;44(2):245–256. doi:10.1016/j.ogc.2017.02.004. Full text

[104] Lo YMD, et al. Presence of fetal DNA in maternal plasma and serum. *Lancet*. 1997;350:485–487. PubMed

[105] Fan HC, Blumenfeld YJ, Chitkara U, Hudgins L, Quake SR. Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. *Proceedings of the National Academy of Sciences*. 2008;105:16266–16271. PNAS

[106] FDA–NIH Biomarker Working Group. *BEST (Biomarkers, EndpointS, and other Tools) Resource*. Silver Spring, MD: Food and Drug Administration; Bethesda, MD: National Institutes of Health; 2016–. Working-group page updated 7 January 2025. Accessed 19 September 2026. NCBI Bookshelf

[107] Thygesen K, Alpert JS, Jaffe AS, et al. Fourth universal definition of myocardial infarction (2018). *Journal of the American College of Cardiology*. 2018;72(18):2231–2264. doi:10.1016/j.jacc.2018.08.1038. DOI; PubMed

[108] Diaz LA Jr, Bardelli A. Liquid biopsies: genotyping circulating tumor DNA. *Journal of Clinical Oncology*. 2014;32(6):579–586. doi:10.1200/JCO.2012.45.2011. Full text

[109] Clark LC Jr, Lyons C. Electrode systems for continuous monitoring in cardiovascular surgery. *Annals of the New York Academy of Sciences*. 1962;102:29–45. doi:10.1111/j.1749-6632.1962.tb13623.x. DOI

[110] Koczula KM, Gallotta A. Lateral flow assays. *Essays in Biochemistry*. 2016;60:111–120. Full text

[111] Luppa PB, Müller C, Schlichtiger A, Schlebusch H. Point-of-care testing (POCT): current techniques and future perspectives. *TrAC Trends in Analytical Chemistry*. 2011;30(6):887–898. doi:10.1016/j.trac.2011.01.019. Full text

[112] World Health Organization. *First WHO model list of essential in vitro diagnostics*. WHO Technical Report Series 1017. Geneva: WHO; 2019. ISBN 978-92-4-121026-3. The list was introduced in May 2018. WHO report; WHO EDL overview

[113] Lassen HCA. A preliminary report on the 1952 epidemic of poliomyelitis in Copenhagen with special reference to the treatment of acute respiratory insufficiency. *Lancet*. 1953;1(6749):37–41. PubMed

[114] Severinghaus JW, Astrup PB. History of blood gas analysis. IV. Leland Clark's oxygen electrode. *Journal of Clinical Monitoring*. 1986;2(2):125–139. doi:10.1007/BF01637680. PubMed

[115] Severinghaus JW, Honda Y. History of blood gas analysis. VII. Pulse oximetry. *Journal of Clinical Monitoring*. 1987;3(2):135–138. PubMed

[116] Kodali BS. Capnography outside the operating rooms. *Anesthesiology*. 2013;118(1):192–201. PubMed

[117] Sjoding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial bias in pulse oximetry measurement. *New England Journal of Medicine*. 2020;383(25):2477–2478. doi:10.1056/NEJMc2029240. Correction: 2021;385:2496. doi:10.1056/NEJMx210003. Full text; Correction

[118] Kennedy HL. The history, science, and innovation of Holter technology. *Annals of Noninvasive Electrocardiology*. 2006;11(1):85–94. doi:10.1111/j.1542-474X.2006.00067.x. Full text

[119] Smuck M, Odonkor CA, Wilt JK, Schmidt N, Swiernik MA. The emerging clinical role of wearables: factors for successful implementation in healthcare. *npj Digital Medicine*. 2021;4(1):45. doi:10.1038/s41746-021-00418-3. Full text

[120] Weed LL. Medical records that guide and teach. *New England Journal of Medicine*. 1968;278:593–600. PubMed

[121] Regenstrief Institute. Get started with LOINC. LOINC. Last updated 2 July 2026. Accessed 1 August 2026. Official site

[122] HL7 International. FHIR overview, version 5.0.0. Accessed 19 September 2026. FHIR R5 specification

[123] World Health Organization. International Classification of Diseases (ICD). Accessed 1 August 2026. WHO

[124] Snow J. *On the Mode of Communication of Cholera*. 2nd ed. London: John Churchill; 1855. Accessed 1 August 2026. Wellcome Collection

[125] National Heart, Lung, and Blood Institute. Framingham Heart Study. Accessed 1 August 2026. NHLBI

[126] Wilson JMG, Jungner G. *Principles and Practice of Screening for Disease*. WHO Public Health Papers No. 34; 1968. Accessed 19 September 2026. WHO IRIS

[127] World Health Organization. Diagnostics. Accessed 1 August 2026. WHO health topic

[128] Yerushalmy J. Statistical problems in assessing methods of medical diagnosis, with special reference to X-ray techniques. *Public Health Reports*. 1947;62(40):1432–1449. doi:10.2307/4586294. DOI; PubMed

[129] Ledley RS, Lusted LB. Reasoning foundations of medical diagnosis: symbolic logic, probability, and value theory aid our understanding of how physicians reason. *Science*. 1959;130(3366):9–21. doi:10.1126/science.130.3366.9. PubMed

[130] Fagan TJ. Letter: Nomogram for Bayes's theorem. *New England Journal of Medicine*. 1975;293(5):257. doi:10.1056/NEJM197507312930513. PubMed

[131] Bossuyt PM, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527. PubMed

[132] Whiting PF, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*. 2011;155:529–536. PubMed

[133] Levey S, Jennings ER. The use of control charts in the clinical laboratory. *American Journal of Clinical Pathology*. 1950;20:1059–1066. doi:10.1093/ajcp/20.11_ts.1059. DOI

[134] Westgard JO, Barry PL, Hunt MR, Groth T. A multi-rule Shewhart chart for quality control in clinical chemistry. *Clinical Chemistry*. 1981;27:493–501. PubMed

[135] National Academies of Sciences, Engineering, and Medicine. *Improving Diagnosis in Health Care*. National Academies Press; 2015. doi:10.17226/21794. Full report

[136] Shortliffe EH. *Computer-Based Medical Consultations: MYCIN*. Elsevier; 1976. See also Buchanan BG, Shortliffe EH, editors. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison-Wesley; 1984. Author's archive at Columbia University DBMI

[137] Doi K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized Medical Imaging and Graphics*. 2007;31:198–211. PubMed

[138] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115–118. doi:10.1038/nature21056. Corrigendum: *Nature*. 2017;546:686. doi:10.1038/nature22985. Article; Corrigendum

[139] Gulshan V, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316:2402–2410. doi:10.1001/jama.2016.17216. DOI

[140] Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *npj Digital Medicine*. 2018;1:39. doi:10.1038/s41746-018-0040-6. Full text

[141] U.S. Food and Drug Administration. Artificial intelligence-enabled medical devices. Updated 4 September 2026. Accessed 19 September 2026. FDA

[142] World Health Organization. *Ethics and Governance of Artificial Intelligence for Health*. 2021. WHO

[143] Best M, Neuhauser D. Ignaz Semmelweis and the birth of infection control. *Quality & Safety in Health Care*. 2004;13(3):233–234. doi:10.1136/qshc.2004.010918. Full text

[144] Fraser DW, et al. Legionnaires' disease: description of an epidemic of pneumonia. *New England Journal of Medicine*. 1977;297:1189–1197. PubMed

[145] McDade JE, Shepard CC, Fraser DW, Tsai TR, Redus MA, Dowdle WR. Legionnaires' disease: isolation of a bacterium and demonstration of its role in other respiratory disease. *New England Journal of Medicine*. 1977;297(22):1197–1203. PubMed

[146] Marshall BJ, Warren JR. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet*. 1984;1(8390):1311–1315. doi:10.1016/S0140-6736(84)91816-6. DOI

[147] Nobel Prize Outreach. The Nobel Prize in Physiology or Medicine 2005: Barry J. Marshall and J. Robin Warren. Accessed 19 September 2026. NobelPrize.org

[148] Corman VM, Landt O, Kaiser M, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance*. 2020;25(3):2000045. doi:10.2807/1560-7917.ES.2020.25.3.2000045. Full text

[149] Vostokov D. *Theoretical Software Diagnostics: Collected Articles*. 4th ed. OpenTask; 2024. ISBN 978-1-912636-91-4.

[150] Wassermann A, Neisser A, Bruck C. Eine serodiagnostische Reaktion bei Syphilis. *Deutsche Medizinische Wochenschrift*. 1906;32(19):745–746. doi:10.1055/s-0028-1142018. DOI

Figure 1. An illustrative montage of diagnostic instruments and evidence, from bedside observation to molecular and computational methods.