Medical Laboratory Informatics

Liron Pantanowitz, MD\textsuperscript{a,*,} Walter H. Henricks, MD\textsuperscript{b}, Bruce A. Beckwith, MD\textsuperscript{c}

\textsuperscript{a}Department of Pathology, Baystate Medical Center, Tufts University School of Medicine, 759 Chestnut Street, Springfield, MA 01199, USA
\textsuperscript{b}Center for Pathology Informatics, Pathology and Laboratory Medicine, Cleveland Clinic Foundation, 9500 Euclid Avenue, L21, Cleveland, OH 44195, USA
\textsuperscript{c}Department of Pathology, North Shore Medical Center, 81 Highland Avenue, Salem, MA 01970, USA

Providing information in a manner that is most effective for patient care is the primary mission of the pathology laboratory. Laboratories are critically dependent on computer systems to manage information. Informatics encompasses the acquisition, organization, validation, storage, retrieval, integration, analysis, communication, and presentation of information, using computers and information technology as tools\textsuperscript{[1,2]}. Information is structured data that can be synthesized into knowledge\textsuperscript{[3]}. Pathology is a data-intensive discipline\textsuperscript{[4]}, and in addition to patient care, data from laboratories are used for documentation of quality assurance, performance improvement, outcomes studies, and research. One common estimate is that 70\% of the objective information used in the management of patients comes from pathology laboratories\textsuperscript{[5]}. To be useful, laboratory information must be accurate, understandable, timely, and available to the correct persons and locations when needed.

Pathology informatics is critical in the profession’s ability to meet its current and future challenges\textsuperscript{[6]}. These challenges include growing laboratory workload with outreach programs; increased adoption of electronic medical records (EMR); and the resulting need to integrate disparate information systems, the shortage of laboratory technologists, patient safety, cost containment, subspecialty centralization, increased demand for molecular testing, and personalized medicine. Many of these challenges can be met by leveraging existing and advancing technologies\textsuperscript{[7]}, such as integrated...
information systems, the Internet, automation, specimen tracking, autoverification, middleware, telepathology, and formal informatics training in pathology residency programs [8,9]. This article provides an overview of laboratory informatics and addresses the core principles and skills required for practicing laboratory informaticists [9].

Computing fundamentals

Computers are composed of hardware (physical equipment) and software. Categories of computers include mainframes (large numbers of users); minicomputers (eg, servers); microcomputers including desktop personal (single-user) computers or portable laptops (notebooks); and hand-held computers like personal digital assistants. Servers, like mainframe computers, share their resources with other computers and simultaneously support multiple users. A computer uses an integrated electronic circuit (microchip) known as a “central processing unit” to interpret computer program instructions, and to process data. The central processing unit exchanges data by a bus (electrical conduit) with subsystems including the basic input-output system; video display controller (connected to a monitor); peripheral or input-output devices (eg, keyboard, mouse, printer); and network interface card. Data that a computer uses are organized as files and are stored on various media that include magnetic hard disks; optical compact disks and digital video disks; flash drives (eg, universal serial bus drives); and magnetic tape. These types of data storage are distinguished from computer memory, which are chip sets that hold programs and data for rapid access by the central processing unit. Two main types of memory are read-only memory and random access memory. Read-only memory is permanent storage of basic input-output system programs that are used for system startup (boot). Random access memory holds programs and data that are in active use and is volatile memory (data are lost when the power goes off or system crashes).

Computer systems recognize and process all data ultimately as a series of 1s and 0s (bits) of a binary number system, which correspond to electronic states of “on” and “off.” Bytes are comprised of 8 bits and are the fundamental unit of data processed by computers. Computer files are collections of data identified by a specific name and purpose. Files are software programs (executables) or data files (text documents, images, audio files). Software refers to computer programs, which are a series of instructions for the central processing unit to execute. The operating system is the set of programs that directs a computer’s functions and file manipulations; examples include Windows Vista and XP, UNIX, MacOS, and Linux. Application software programs direct the computer to perform a specific function (ie, tailored to a specific “application” of the computer’s resources and capabilities). Application examples include word processors, spreadsheets, World Wide Web browsers, and laboratory information systems (LISs).
To leverage the time and expertise of a pathologist or laboratory scientist, a comprehensive workstation is required [10]. Components of a successful paperless workstation (Fig. 1) include access to the LIS and Internet; multimedia capabilities; and dedicated applications depending on the specialized needs of the laboratorian (eg, statistical analysis, image analysis).

Networks

A network refers to computers and devices connected by a telecommunications link made using cables (copper, fiberoptic) or electromagnetic waves (wireless). Computers and other devices connect to networks by network interface cards. Network topologies include star (central hub); token ring (sequential connections); linear bus (central backbone); and tree (branched) networks [11]. Software (network operating system) is an essential component of any network. A local area network links multiple devices within a small geographic area (eg, hospital). Wide area networks are used for broader geographic areas. Virtual private networks are a widely used method of securely sending confidential information across a public network, such as

Fig. 1. Surgical pathology workstation. Components of a successful paperless workstation include (1) easy access to a multimedia personal computer with significant computing power connected to the LIS, EMR, and Internet; (2) digital camera; (3) voice activation; and (4) split keyboard.
the Internet. Wireless networks, although convenient for mobile users, are subject to interference and require special attention to security.

Complex information systems, such as LISs, function in networked environments. Typical architectures include mainframe, client-server, and thin client. In the mainframe model, all software functions, transactions, and data reside on a single large computer. Users access the system over a network by “dumb” terminals that have no functions other than data input and display. In a client-server system, a client refers to a computer that accesses a remote service on a network server. One or more servers may be present that provide specific functions (eg, fax server). Software functions are distributed across client and servers as opposed to the centralization on a mainframe. Clients can be classified as thick (fat), thin (lean), or hybrid clients. A thick (fat) client (eg, workstation) independently performs the bulk of any data processing operations itself. A thin (lean) client (eg, terminal), however, depends primarily on a central server for processing activities, and focuses largely on the input and output of data. A hybrid client is a mixture of these two. Thin clients have lower costs for hardware and can better leverage existing hardware, although there are licensing fees for thin client set ups. Thin clients are also easier to secure and update, because administration occurs centrally. Thin clients also demand less network bandwidth, because terminal servers typically reside on the same high-speed network backbone as file servers, and most network traffic is confined to the server room. In a thin client environment, only mouse movements, keystrokes, and screen updates are transmitted from and to the end user, unlike large files or documents with a thick client.

The Internet is the global network of computer networks that includes international telecommunications infrastructure. Internet functions include e-mail; World Wide Web (web or www); and file transfer. An intranet is a private network based on the same network protocols as the Internet. The World Wide Web is a conglomeration of documents accessible by a particular communication protocol over the Internet. Hypertext markup language is used to build web pages and includes characters, graphics, and hyperlinks. Hypertext is a method for displaying web pages in programs known as “web browsers.” Hyperlinks in web pages enable navigation (surfing) to other web pages. Information on the World Wide Web has become increasingly useful to pathologists [12–15], and has provided new opportunity for training and education [16]. World Wide Web–based technology has offered physicians, and even patients, rapid access to laboratory test results. The Internet has also provided a mechanism for rural or underserved areas to gain access to health care.

Digital data can be transmitted by (1) serial (1 bit at a time) or parallel (data moves simultaneously over multiple wires) communication; as (2) synchronous, isosynchronous, and asynchronous communication; and (3) simplex (unidirectional), half-duplex (bidirectional, but one direction at a time), and full-duplex (bidirectional, both directions simultaneously) communication [17]. Bandwidth (bits per second) refers to the amount of information
that can be sent over a network connection in a given period of time. Transmission rates vary according to the network connection [18]: 56 kilobits per second for “plain old telephone service” using a dial-up modem, 1.5 megabits per second for a T1 line, 128 kilobits per second to 24 megabits per second for digital subscriber line, up to 30 megabits per second for a cable modem, 45 megabits per second for a T3 (DS3) line, 54 megabits per second for 802.11g wireless (wi-fi), and Ethernet cable up to 100 megabits per second. Several protocols (rules or algorithm) are available that describe how data are transmitted over a network. Protocols define the syntax (data structure or format), semantics (data interpretation), and synchronization of exchanged data. Common protocols are transmission control protocol, Internet protocol, hypertext transfer protocol, and file transfer protocol.

Databases

Databases are large collections of data organized into fields (columns in a table); records (rows); and files [19]. They provide a flexible and organized way to store, retrieve, and manipulate data to create information, and are the preferred method of storage for large multiuser applications like the LIS. Large integrated databases (repositories and warehouses) can be analyzed (data mining) to identify patterns or relationships [20]. Data repositories integrate diverse data from multiple systems, often form part of the hospital information system (HIS), and underpin the EMR. Data warehouses are also large integrated databases, but are better structured to support efficient queries. Because warehouses do not support large numbers of small transactions well, they usually do not connect to clinical systems [21]. The LIS is at its core a database that determines the configuration of system parameters and that stores patient-related data.

Database models include flat-file (single two-dimensional table); relational (multiple related tables); hierarchical (data organized into a tree-like structure); and object-oriented databases. Computer programs (database engines) used to manage and query (retrieve) databases are known as “database management systems” (eg, Oracle, Microsoft Access, Sybase). Data are usually retrieved by using a query language. Structured Query Language (SQL) is a computer language used to create, retrieve, update, and delete data from relational database management systems. Open database connectivity is a standard or open application programming interface for connecting different database management systems. This application programming interface is independent of any one programming language, database system, or operating system [2].

Laboratory information systems

The LIS is at the core of most pathology laboratory operations. Its functions include workflow management, specimen tracking, data entry and
reporting, assistance with regulatory compliance, code capture, interfacing with other systems, archiving, inventory control, and providing billing information [22,23]. It can also be used for quality assurance measures [24]. Components of the LIS include hardware (e.g., servers); peripherals (e.g., instruments, printers); a network; interfaces to other information systems; databases; and software, such as an operating system, database management systems, and specific applications required for laboratory operations. A LIS can be a stand-alone (best-of-breed) system or form part of an enterprise (hospital-wide) integrated HIS (single vendor solution) [25].

Within the laboratory, there is also a choice between integrated or separate systems for anatomic pathology, clinical pathology, and the blood bank. Integrated systems have one vendor, database, set of dictionaries, log-on, report type, and invoice. They also eliminate the need to interface data between different laboratory departments. The ability of the LIS to support specialized areas of the laboratory, such as molecular diagnostics, cytogenetics, flow cytometry, and tissue typing, is becoming increasingly important. These specialty laboratories do not conform to current LIS models, and they represent a new challenge for information management in pathology informatics [26].

Regardless of the type of system, it is important to keep current with software updates, because outdated versions are often sunsetted (i.e., no longer supported by vendors). Laboratories usually purchase LIS hardware and software and physically house the system in the organization. As an alternative, in the application service provider model the laboratory rents an entire remotely located web-enabled LIS (hardware, software, and data storage) from the application service provider. A web-enabled LIS refers to a LIS or specific application accessed by the Internet, or one that delivers access to others by the Internet (e.g., web-based outreach) [27]. Increasingly, more vendors offer physician office-laboratory link software to provide web-based connectivity (electronic outreach). These laboratory Internet portals not only provide secure Internet connectivity to their clients, but also improve the marketability of laboratory services. They allow remote ordering, test catalogs, results inquiry, and reporting at any location or to any device, including hand-helds with Internet access.

**Dictionaries and worksheets**

LIS dictionaries (Fig. 2) and worksheets define the conventions and logical framework for information processing and workflow throughout the laboratory. LIS dictionaries (maintenance tables) are database tables or files that store and maintain information used repeatedly during activities, such as accessioning of specimens. They help standardize and structure protocols; procedures; terminology and codes (e.g., billing codes); control workflow; provide security features; define rules and limit selections for data fields; improve entry of valid data; define report content and format elements; and
automate billing. Dictionaries provide choices in look-up windows or drop-down lists. Examples include worklists (worksheets); logs; autoverification parameters; standardized nomenclature; and security and access level privileges for users. Dictionaries may be prebuilt by the vendor, but typically the laboratory defines and maintains the entries in a manner that meets its specific needs.

LIS worksheets (or worklists) define and accordingly group tests to be performed at a certain laboratory location, workstation, or instrument. Worksheets organize laboratory workflow. In anatomic pathology LISs, logs are analogous to worksheets and define the workload for a given area or time period. For instance, a histology log indicates what tissue blocks to cut and which stains to perform.

Interfaces

An interface is a hardware and software link that connects two computer systems, or a computer and its peripherals, for data communication. Interfaces rely on standards that enable incompatible systems to exchange data (see later section on Standards). LIS interfaces (Fig. 3) include instrument-analyzer interfaces and application interfaces with other systems. Most tests in a clinical laboratory are performed on automated instruments that exchange data directly with the LIS through unidirectional or bidirectional interfaces. Common application interfaces with the LIS include EMR, admission-discharge-transfer, and financial systems. An admission-discharge-transfer interface obviates the need for manual entry of patient demographic data directly into the LIS. Most laboratories report results electronically through an interface to an EMR or HIS, and the LIS is
increasingly being called on to exchange data with other information systems [28]; web portal servers; tumor registries; practice management systems; and devices (eg, automated stainers). Such interfaces can be leveraged to improve efficiency and automation, and eliminate potential sources of error [29]. Laboratories may also receive test orders interfaced from these systems.

**Workflow**

The LIS supports workflow and information flow in all steps (preanalytic, analytic, and postanalytic) of the laboratory testing process (ie, it closes the loop between the test order and result delivery) [1]. The preanalytic phase involves patient registration (if the order is not received from an external system); test order and selection; specimen collection and labeling (frequently LIS-generated); specimen receipt (accession number assignment); and tracking. The analytic phase includes work distribution (worklists) and specimen preparation (eg, aliquots and bar code labeling); test performance and analysis; test interpretation; possible additional testing (eg, reflex testing, immunohistochemistry); result entry (eg, interfaced, manual, transcription); and verification (manual, automatic release, electronic signature). In the postanalytic phase, the LIS allows the generation and delivery of laboratory reports; test results (printing, faxing, electronic transmission); and modification of reports (amendments and addenda).

Certain areas of the laboratory have specialized LIS needs [1,30]. Microbiology data are more complex than many other areas, with the need to
handle preliminary culture reports and updates, antimicrobial sensitivities, and epidemiology reporting. In the blood bank, the LIS needs to track blood components and their derivatives, maintain inventories of blood products and donors, perform safety checks between patients and blood products from order entry until transfusion, and retain patient immunohematologic or special needs histories for prolonged periods. For point-of-care testing, the LIS needs to be able to receive data, entered directly into the LIS or uploaded by an interface from different types of wired or wireless devices. To meet regulatory requirements, cytopathology needs to exploit databases within the LIS to perform quality assurance measures, such as cytologic-histologic correlation.

**Reporting**

The LIS can print a wide variety of patient reports including interim, cumulative, discharge, and order-based report types. Laboratory reports (paper and potentially electronic) are required to include (when appropriate) the following data elements: unique patient identification (e.g., date of birth); name and address of the performing laboratory; report date; tests performed; specimen source; result; units of measure; reference range as determined by the laboratory performing the test; and information regarding specimens that do not meet acceptability criteria [31,32]. Formatted reports (e.g., font, tables) are often not accommodated by particular interfaces or by the display in an EMR. Synoptic (structured or formatted) reporting in surgical pathology, in which the pathologist completes prearranged data entry templates (checklists), allows these data elements to be stored in a relational table within the LIS, provides clinicians and cancer registries with this information in a standardized manner, and facilitates data extraction [33,34]. Some vendors have included such checklists into their LIS, or offer them as a stand-alone system that can be integrated with the LIS.

**Rules**

A major advantage of the LIS is that it can easily perform calculations and execute algorithms or rules. This improves productivity, reduces staff needs, improves consistency, reduces errors, and speeds up workflow [35,36]. An example of a simple algorithm is a delta check, in which a patient’s current result is compared with their prior result, and the result is flagged if the difference exceeds a specified limit. The LIS can also be programmed to flag abnormal results for review or autoverification (see below). Some blood banks have moved to an electronic (computer-assisted) crossmatch to confirm ABO compatibility between patient and donor, without performing a serologic crossmatch [37]. The electronic crossmatch has been successfully performed even in sites remote from the blood bank, such as the operating room (virtual blood banking) [38].
Autoverification

Autoverification (autovalidation) refers to the automatic release of results received from a laboratory instrument without technologist review. A test result transmitted over an instrument interface is evaluated for “pass” or “fail” by the LIS, based on parameters that the laboratory defines in the system. If it passes, then the result is automatically released by the system (i.e., no manual review is required). Any value that fails (i.e., falls outside defined parameters) requires manual review by a designated operator (e.g., laboratory technologist). Customized rules used for autoverification can be simple or complex. They may include reference ranges, checks that quality control was passed, critical values, delta checks, dilution needs, instrument flags, laboratory review policies, and specific patient locations or ordering physicians. Autoverification can be enabled through the laboratory instrument, middleware, or LIS. For laboratories that want to autoverify using results originating from multiple instruments, autoverification needs to be performed using middleware, or the LIS. Benefits of autoverification include consistency of applying decision rules across all shifts at all times, decreased turnaround time, better use of staff, error reduction, and improved patient care [39].

Middleware

Middleware (formerly referred to as “interface devices” or “data managers”) is any software tool placed between a laboratory analyzer and the LIS that augments the LIS capabilities. Middleware may be a cost-effective way to add functionality to an older (legacy) LIS. In addition to facilitating interface implementation, middleware has allowed laboratories to improve result-handling mechanisms for autoverification, reflexive testing algorithms, outreach systems, point-of-care testing, image management, and tracking specimen storage [40].

Administration

Administrative aspects related to the LIS include purchasing a LIS, implementation, validation, change management, regulatory compliance, licensing, and security. Purchasing a LIS involves planning, usually by a committee of stakeholders; preparing a request for information; conducting a gap analysis (analyzing current and new LIS requirements); cost analysis; determination of workflow impact and licenses needed; development of system specifications; development and distribution of a request for proposal; onsite demonstration; site visits; system selection; and contract negotiation [41–43]. A request for proposal is a document supplied to the vendors that details all required system functionality including functional, technical, training, and implementation requirements.
System implementation includes all the tasks that are necessary to get the LIS installed and operating. Steps include dictionary building, system configuration, testing, historical data conversion, startup, interface validation, documentation, and training. One of the most important and most time-consuming parts of LIS implementation is configuring the dictionary tables and building worksheets. Dictionaries are typically built in a particular sequence, because some table entries are based on choices from pre-existing tables. A thorough, careful approach to this phase of implementation is key to successful adoption of a new LIS. Any new (or upgraded) LIS needs to be validated to ensure that it performs in the way intended. Changes need to be migrated from a test (development) environment to live (production) environment through a change control procedure. Validation of the LIS needs to be a continuous process, all computer systems used within the laboratory need to be maintained, and software kept current [44]. This should be documented and readily available for regulatory and accreditation entities, such as the College of American Pathologists, the Joint Commission, American Association of Blood Banks, and others. In the United States, blood bank software is considered a medical device and must be cleared by the Food and Drug Administration [25].

Data security practices and policies are extremely important for protecting the privacy, confidentiality, and integrity of patient data [45]. United States laboratories need to comply with the Health Insurance Portability and Accountability Act. Physical safety of the LIS requires appropriate computer room facilities (controlled humidity and temperature, fire protection, and secure access); uninterruptible power supply; and data backup. Regular data backup is perhaps the single most important measure with respect to data integrity [46]. In case of disasters (eg, fire, flooding), a disaster recovery plan should exist. Because computers and software can malfunction, and networks can disconnect, the laboratory must have a backup plan for operations during both scheduled and unscheduled downtime (unavailability) of the LIS. Computer downtime can be associated with adverse clinical outcomes or additional staff expense [47]. Security for a web-enabled LIS can be achieved by a secure, encrypted network, virtual private network, or other means. Measures to prevent a breach of data security include encrypting transmitted information; authentication (eg, access codes, passwords, key cards, biometrics); firewalls; malicious intrusion protection (eg, antivirus software); and audit trails (computer-generated, time-stamped electronic records that reconstruct the course of events relating to the creation, modification, and deletion of an electronic record).

Standards

The aggregate of rules, formats, and functions for seamlessly transmitting data between components in a system or network is called a protocol (standard or specification). Standards define how to encode identifiable data and
how to package and communicate this information [17]. Standards development organizations include the American National Standards Institute, International Standards Organization, and American Society for Testing and Materials (ASTM). Examples of relevant standards include

- **Health Level Seven (HL7)** is currently the most widely used application standard for data exchange in health care. HL7 messages consist of segments, each comprised of fields (Fig. 4). Translation tables may be needed to cross-reference different HL7 codes specified by different vendor systems. Several versions exist (eg, HL7 v2.x, v3.0). The HL7 standard uses an information model (reference information model) to represent real-world objects and concepts. In addition, they promote the clinical document architecture, which is an extensible markup language (XML) markup standard that specifies the structure and semantics of “clinical documents” for the purpose of exchange.

- **ASTM** is a commonly used protocol for sending laboratory data between the LIS and analyzers.

- **Transmission control protocol–Internet protocol** is the suite of communications protocols used to connect computers, systems, and network devices on the Internet and internal networks.

- **Digital Imaging and Communications in Medicine (DICOM)** is a standard for handling, archiving, printing, and transmitting medical imaging information. It includes file format definitions and a network communications protocol. DICOM has been widely adopted by hospitals, principally for radiology imaging. DICOM extensions designed to accommodate pathology have been adopted [1].

- The **World Wide Web Consortium** created XML, a general purpose markup language intended to facilitate data exchange among different information systems. XML documents can be the basis for a hypertext markup language document (web page), a PDF document, and even a Microsoft Word document from the same master file. The purpose of XML is to structure data and describe documents in a way that facilitates exchange and analysis. XML representations of pathology reports

```
0001 MSH|^~&|HIS|BA6|APLIS||200705041500|SMS
0002 EVN|A04|200705041500
0003 PID||0123|5432^HOSPITAL||NAME||F||W|POBOX||TEL|ETC
0004 PV|0001|OP|RADIOLOGY^^|R|||DOCTOR^RAD||ETC
0005 DG|0001|||HISTO||
0006 GT|0001||NAME||POBOX|TEL||SELFPAY|ETC
0007 IN||0678945|UNKNOWN
```

Fig. 4. Example of an abbreviated inbound HL7 formatted message from the HIS to the Anatomical Pathology LIS (APLIS). The message contains seven segments including a message header (MSH); event details (EVN) about what triggered the message; the patient’s identification (PID); and information about this particular patient visit (PV), the diagnosis (DG), and their guarantor (GT) and insurance (IN) data for billing purposes.
are a promising standard format. With XML, pathologists can annotate all of their data in a format that can transform every pathology report into a database, without compromising narrative structure [48].

Digital imaging

A digital image is represented in a computer by a two-dimensional array of numbers (bitmap or raster image), each element of which represents a small square area of the picture, called a picture element or pixel [49]. The bit depth refers to the number of colors available (eg, 24-bit image has 16.7 million colors). Image resolution (megapixels) depends on the overall (width \times height) number of pixels (eg, 640 \times 480 image = 0.31 megapixels), whereas the file size is the bit depth multiplied by image resolution (eg, 24-bit \times 0.31-megapixel image = 900 kilobytes). Image file formats include joint photographic experts group (jpeg), joint photographic experts group 2000, graphic interchange format (gif), tagged image file format (tiff), bitmap, and portable network graphics (png). Images can be transmitted or stored in a compressed form (reduced image size), achieved by removing redundant information. Compression algorithms may be “lossless” (no loss of data) or “lossy” (some detail is lost). Digital images can be created by a variety of input devices (eg, digital cameras, document scanners) or synthesized (eg, computer graphics). Digital cameras use charge-coupled devices or complementary metal oxide semiconductor image sensors to measure light energy, and added circuitry to convert the measured information into a digital signal. Because digital cameras do not use photographic film, images are immediately available for viewing and do not degrade over time.

The imaging process entails four steps: (1) capture; (2) saving (storage); (3) editing if necessary; and (4) using (viewing, displaying, printing). This process as it relates to applications in pathology has yet to be standardized [50]. Good photographs depend on several factors including the camera lens (eg, macro capabilities for gross photography); microscope optimization (eg, Koehler illumination); white balancing; flatfield correction; global and focal adjustment [51]; image editing (eg, sharpening); and compression. If image manipulation is performed this should be documented, and consideration should be given to storing both the original (raw) and edited (modified) image [51]. Because digital images are easily manipulated [52], the potential fraudulently to alter them has raised much concern [53,54]. As a result, forensic-type software is emerging to trace tweaks to digital images [55]. In a clinical laboratory that captures many digital images, scalable image storage requirements are needed. Total storage needed can be estimated as (number of image files) \times (average file size) \times (1.25). Virtual microscope slides (see below) require an extremely large amount of server storage if they are used for routine practice. From experience, implementation of digital imaging in a pathology department often leads to an increase in the number of images taken [56,57], perhaps related to the ease
of taking digital pictures. Image management systems of digital images are essential for the clinical use of pathology images and as DICOM use expands, laboratory images will be managed by software analogous to the picture archive and communication system currently used by radiology [58].

Pathologists are increasingly integrating digital images into their practice of medicine [59]. Digital images can be used for diagnosis; communication; quality assurance studies (eg, peer review); computer-assisted image analysis (eg, multispectral imaging); automated screening of Pap test slides; proficiency testing; archiving; research; publication; education; and training [60–64]. Image-enhanced reports are a growing trend among pathology practices [65]. LIS vendors have begun integrating digital image acquisition and storage modules into their products.

The emergence of technology that supports digital imaging along with greater image quality, and higher processing capacity of computers, has promoted the use of telepathology [66]. Telepathology (including teletypology) is the practice of pathology at a distance by using telecommunication to transmit images. Telepathology can be used for diagnosis, consultation, or education. Earlier studies found the accuracy of telepathology to be less than that of light microscopy [67]. With technologic advances, however, more recent studies have demonstrated improved accuracy and reproducibility [68], even despite image compression [69]. There are three types of telepathology systems: (1) static; (2) real-time (dynamic); and (3) virtual (whole) slide imaging systems. Static image systems are cheaper, but can only capture a selected subset of microscopic fields. The latter two systems permit evaluation of the entire slide, but are more costly, and may be hampered by high network traffic. With some real-time telepathology systems, the consultant can actively operate a remote microscope with a robotic stage. Validation (of both system and user), reimbursement, and medicolegal issues surrounding telepathology need to be refined [70,71].

Virtual microscopy and whole-slide imaging, the use of digital imaging to produce digital (histology and cytology) slides that simulate light microscopy, are being used not only for telepathology, but also for archiving, clinical diagnosis, and education [72]. Virtual microscopy provides access to all areas of interest on a slide by using a personal computer or digital device, without the use of a microscope. Systems are now capable of complete digitization of slides at high magnifications, a process known as “whole-slide imaging” [73]. Selected scanning systems can digitize multiple focal planes, to create a virtual slide with the ability to focus. Virtual slides have been introduced into the certification examination for the American Board of Pathology, and may soon provide an effective tool for proficiency testing [74]. Current automated high-speed whole-slide imaging systems are sufficient for diagnostic purposes and potentially represent a disruptive technology in the traditional practice of pathology.
Coding

Coding is the process of classifying data, by assigning a representation within a predefined taxonomy. Coding provides a predictable, consistent, reproducible, and structured (standard) language for a set of concepts. Coding normalizes concepts, permits abstraction from documents or files, allows for automated data processing, facilitates communication, and permits relationships between concepts to be highlighted. Many different types of health care data can be represented using a variety of coding systems [75]. Coded data are used to transmit and retrieve information for patient care, reimbursement, performance improvement, planning, facility management, and research. The laboratory informaticist should ensure that official coding conventions are followed and that codes are accurately assigned. Outside agencies and payors often monitor the accuracy of coding because of its impact on reimbursement. Coding can be performed manually or by using automated systems. Coding software (encoders) can use either a branching logic system (code assignment guided by a series of questions) or automated codebook (index or list) approach. Examples of coding systems include

- American Standard Code for Information Interchange (ASCII) assigns a number to each key on the keyboard that can be exchanged and read by most computer systems.
- Current Procedural Terminology (CPT) is used to provide a numeric coding scheme for diagnostic and therapeutic procedures for billing purposes. Laboratorians need to be vigilant about the correct use of these codes because upcoding, downcoding, and unbundling of these billing codes are fraudulent.
- International Classification of Disease (ICD) is a hierarchical listing used for classifying diseases and related symptoms and signs. The most recent revision is the tenth edition. As of 2007, however, most United States payers require the use of International Classification of Disease-9-CM (ninth revision, clinical modification). These codes are commonly used to justify Current Procedural Terminology codes submitted for reimbursement.
- Logical Observation Identifiers, Names, and Codes (LOINC) is a numeric code system aimed at standardizing laboratory and clinical codes for use in clinical care, outcomes management, and research. Logical Observation Identifiers, Names, and Codes (over 25,000) works within HL7 messages to standardize test names and codes. Codes do not exist, however, for all laboratory results, there are many similar codes, and assignment of these codes is difficult.
- Systematized Nomenclature of Medicine (SNOMED) is a standardized medical vocabulary system originally designed for describing pathologic findings in medical databases. It has been extensively expanded and current modules contain more that 144,000 terms, which are available in multiple languages. Systematized Nomenclature of Medicine is widely
used, well suited to the needs of pathology, and is likely to become an international standard.

Hospital information systems

The formation of health care organizations through the consolidation of health care providers into integrated delivery networks was accompanied by the evolution of HIS [76]. The HIS environment includes several upper-level systems, such as the clinical information systems; decision support systems; admission-discharge-transfer system; management information systems; business and financial systems; communications and networking applications between departments; and departmental systems, such as the LIS [76,77]. The HIS is the nexus for inpatient-related activity, and is often the location for the master-patient index. Many hospitals have an interface engine at the hub of their HIS architecture, to manage the flow of data between disparate components of their information systems.

Electronic medical record

The EMR represents the medical-health record of a patient in electronic format. The EMR is intended to provide secure, real-time, current, interactive, patient-centric information to aid clinical decision-making. EMRs aim to provide access on a computer or over a network to a patient’s health information at any point of care. It typically comprises health information from several locations and sources, and includes demographics; medical history; examination and progress reports of health or illnesses (continuity of care record); scheduling; laboratory and other test results; possible image display (eg, radiographs); medications; clinical practice guidelines; and billing information. Laboratory transactions often comprise the largest portion of information that goes into the EMR. Because the EMR is the portal through which many clinicians now order laboratory tests and interact with laboratory information, it has several implications for the laboratory [78]. The EMR provides access to clinical information that pathologists may use to integrate into their interpretations, and allows data mining opportunities for quality and outcome analyses. Most EMRs are currently supported by medical (nonpathology) informaticians. Pathologists and laboratories are at risk of losing influence over the manner in which laboratory information is handled in the EMR. One role of the laboratory informatician is to ensure that EMR users reading electronic laboratory reports do not misinterpret the data because of the manner in which it is reported [79].

Laboratory data from the LIS are typically transmitted to the EMR by an HL7 interface. The format and display of these results, and any
accompanying information (e.g., interpretive comments), are dictated by the screen design in the EMR. Corrected reports transmitted to the EMR need to replace (overlay) previous results. At present, pathology images are not typically available in the EMR, although this may change as more LIS vendors adopt DICOM or other imaging solutions. Particularly relevant to the clinical laboratory, computerized physician order entry (CPOE) refers to the electronic entry of physician test orders. CPOE reduces errors related to handwriting or transcription, and incomplete clinical information on paper requisitions [80]. It also provides opportunities for pathologists to influence test ordering patterns through structured order screens, and decision support rules and alerts, which trigger at the time of order entry. Moreover, CPOE permits analysis of real-time data to assess the impact of changes, something not possible with paper-based systems [81]. For CPOE to be beneficial to the laboratory, however, laboratorians need to be active participants in planning the implementation of CPOE in their EMR [82]. Currently, there are very few standards specific to the EMR, such as the Australian [83] and ASTM [84] standards for an electronic health record architecture. Standardization is required to bring uniformity in EMR development, and to address connectivity, security, and confidentiality of health care data.

**Future trends**

Several transformative technologies along with emerging clinical, economic, and research demands on the practice of laboratory medicine necessitate informatics participation. These include miniaturization (moving testing from the laboratory to the bedside); radiofrequency identification; electronic document management systems; voice transcription; electronic transmission of pathology data to tumor registries; biospecimen repository oversight; and virtual pathology [85]. Virtual pathology is the practice of diagnostic pathology whereby the technical performance of an analytical technique (e.g., flow cytometry) is performed at one location and the necessary elements are transmitted in electronic form to another site for diagnostic interpretation [86]. The advent of more sophisticated techniques and algorithms for genomic and proteomic testing in pursuit of personalized medicine promises to result in data complexity and volume that greatly tax or overwhelm capabilities of current systems. The development of more robust pathology information systems that facilitate data integration from diverse modalities, data presentation, and decision support is essential to support pathologists in the twenty-first century [6]. Furthermore, as health care systems expand and incorporate disparate information systems within their organization, or with other organizations, interoperability and the promotion of standards is key. The laboratory informatician needs to play an increased integrative role in their laboratory, health care organization, and region.
References


