Technology Assessment Program

Office of Patient Care Services

UPDATED INFORMATION FOR VA TECHNOLOGY ASSESSMENT PROGRAM (VATAP) REPORTS

In June 2000, VATAP was relocated within the Veterans Health Administration from the Office of Research & Development to the Office of Patient Care Services. The following report was produced prior to the relocation of VATAP.

Current VATAP contact information is as follows:

VA Technology Assessment Program (11T)

VA Boston Healthcare System

150 South Huntington Avenue

Boston, MA 02130

Tel: 617.278.4469 Fax: 617.264.6587
vatap@med.va.gov
http://www.va.gov/vatap http://vaww.va.gov/vatap

Technology Assessment Program

Report No. 5

Picture Archiving and Communication Systems:

A Systematic Review of Published Studies of Diagnostic Accuracy, Radiology Work Processes, Outcomes of Care, and Cost

Authors: *Diana Anderson,R.N., M.P.H., Technology Assessment Program Research Analyst*
Karen Flynn, D.D.S., M.S., Manager, Technology Assessment Program

Contributors: *Elaine Alligood, M.L.S., MDRC Information Center Librarian*

Report Date: August 1997

MDRC
MANAGEMENT DECISION AND RESEARCH CENTER
HSR&D
Health Services Research & Development Service

MTA95-001-01

The Health Services Research and Development Service (HSR&D) is a program within the Veterans Health Administration's Office of Research and Development. HSR&D provides expertise in health services research, a field that examines the effects of organization, financing and management on a wide range of problems in health care delivery, quality of care, access, cost and patient outcomes. Its programs span the continuum of health care research and delivery, from basic research to the dissemination of research results, and ultimately to the application of these findings to clinical, managerial and policy decisions.

Diana Anderson, B.S.N., M.P.H., Research Analyst
Technology Assessment Program
Management Decision and Research Center (152M)
Health Services Research and Development Service
Office of Research and Development
VA Medical Center
150 South Huntington Avenue
Boston, MA 02130

Tel: (617) 278-4469 FTS: (700) 839-4469 Fax: (617) 278-4438
elizabeth.adams@med.va.gov
g.mdrc-ta@forum.va.gov

Released: January 1998

ACKNOWLEDGEMENTS

The contributions of the following reviewers are gratefully acknowledged. The MDRC takes full responsibility for the views expressed herein. While the observations and suggestions of reviewers have been of enormous assistance, participation does not imply endorsement.

M. Anthony Graham, M.D.	Chief of Telemedicine Department of Veterans Affairs
Douglas D. Bradham, Dr.P.H.	Director, HSR&D Developmental Center Baltimore VA Medical Center

The MDRC is most grateful to the following people for their support and contributions to this project:

John R. Feussner, M.D.	Chief Research and Development Officer Department of Veterans Affairs
Thomas V. Holohan, M.D.	Chief Patient Care Services Officer Department of Veterans Affairs
Kathleen Meyers	Chair, Baltimore PACS Steering Committee HOST program, Martinsburg VA Medical Center

The MDRC would like to also wishes acknowledge the contribution of Becky Sarah, M.P.H., for her assistance in the development of this report.

Picture Archiving and Communication Systems: A Systematic Review of Published Studies of Diagnostic Accuracy, Radiology Work Processes, Outcomes of Care, and Cost

PREFACE

> *"The convergence of healthcare technology and telecommunications technology offers an extraordinary opportunity to expand the availability and affordability of modern healthcare. Whether it is long-distance video-conferencing with specialists, the transmission of images or data, the availability of patient information, or medical education materials on the Internet, telemedicine expands access to healthcare" (FCC, 1996).*

The experience of private industry in the last decade suggests that a robust information technology infrastructure is an essential element for managing a complex organization. Many private industries also recognized that, in order to capture the potential benefits of sophisticated information technologies, they needed to deploy them in an environment that supported complex changes in organizational culture and work process re-engineering (Clyburn, 1996).

The veterans health care system is now in a period of transition. System-wide efforts are underway to re-engineer the structure and work processes of the organization, guided by a strategic plan to transform VA into an integrated health care system that provides convenient, responsive high-quality care in a cost-effective manner. Information systems have been described as the "bricks and mortar" which will hold the health care system together in the future (Kizer, 1996). New information technologies, particularly selected applications of telemedicine ("medicine at a distance"), are being implemented to support organizational change. These technologies offer VA the potential to enhance the quality of care, to improve access to care, and to control costs.

Picture Archiving and Communication Systems (PACS) are high-speed, graphical, computer network systems for the storage, retrieval , and display of radiologic images. Within a single organization, they may be installed as a "mini-PACS" in a specialty department (often radiology, the ER, or an ICU) or may be integrated into a hospital-wide information system. With the development of integrated health care systems supported by the appropriate information infrastructures, PACS can be implemented over distant networks to provide remote access to patient information and to support one of the most promising forms of telemedicine, teleradiology.

There are many legislative, licensing, legal, and reimbursement issues related to the use of PACS, teleradiology, and telemedicine that need to be resolved. The federal government, private industry, and members of the national and international scientific and legal communities are working to help resolve these issues. As initiatives continue to upgrade the existing infrastructure and to remove obstacles to implementation, telemedicine applications and related technologies are evolving rapidly. VA continues to play a major role in the development and dissemination of medical information technologies.

VA recognizes the continuing need to evaluate these emerging technologies from a broad range of perspectives, to ensure that veterans continue to receive safe, high quality health care throughout this period of transition. VA has provided support for assessments of a variety of telemedicine and other information technologies. The attached report, a systematic review of the literature regarding the use of Picture Archiving and Communication Systems (PACS) in a clinical setting, was prepared in response to requests by VA's Office of Research and Development and the Hybrid Open Systems Technology Program.

EXECUTIVE SUMMARY

Purpose

Historically, VA has played a major role in the dissemination of information technology into medicine as part of its coordinated efforts to improve the efficiency and effectiveness of its health care delivery system. To support these efforts, the Office of Research & Development of the Department of Veterans Affairs (VA) requested that the Management Decision and Research Center (MDRC) define significant issues and frame potential future research questions related to the use of picture archiving and communication systems (PACS) in clinical settings. This document will also clarify for VA decision makers what is known, and not known, about the impact of PACS when used in a clinical setting.

Methods

The MDRC performed a systematic review of the published evidence about the clinical performance and economics of PACS. A systematic review uses a rigorous scientific approach to limit bias, and provides evidence-based conclusions. This review complements the technical research and development efforts already underway within VA.

Significant Evaluation Questions

We identified published research that addressed four critical questions about the use of PACS in clinical settings:

1. How does the **diagnostic accuracy** of interpreting digital images on a workstation or printed to film (either or both available with PACS) compare with that of conventional film-screen radiology?
2. How does **work process efficiency** in a PACS environment compare with that of a conventional, largely film-based, radiology department?
3. What are the effects of PACS on clinical care and **patient outcomes**?
4. What **cost savings** accrue from changes in processes and outcomes of care that can be attributed to PACS?

Findings

Rapidly changing technologies that are embedded in the organizational infrastructure are inherently challenging to assess. It is particularly difficult when the technology itself is ill-defined. PACS vary widely in their configuration and in their use within organizations. This diversity has contributed to difficulties in developing methods to assess PACS that are broadly applicable to multiple settings, and suggests that many study findings may be difficult to generalize to other settings.

While many published articles express enthusiasm about the potential for PACS and related technologies, few provide scientifically sound evidence of clinical effectiveness, efficiency, and cost-effectiveness. Only 3% of the reviewed literature contained evidence which could be included in this report. The findings of our systematic review are summarized below.

Diagnostic accuracy: safe and accurate functioning needs to be established before studies of clinical and organizational impact become appropriate. Most of the clinical studies of PACS identified in the literature search addressed the issue of diagnostic accuracy.

- published evidence did not consistently show that the diagnostic accuracy of digital imaging (read on a PACS workstation or printed to film) was equal to that of conventional X-ray film.
- recent improvements in digital imaging technology are reflected in research findings.

Work process efficiency: few studies examined work processes in radiology, identified bottlenecks, or reported work process changes related to installing a PACS.

- compared with conventional film management, PACS appeared to save time when generating, retrieving, and delivering images.
- overall, physicians did not obtain the results of imaging exams any sooner with PACS than with film, despite the fact that the information was actually available 50 minutes sooner with PACS than with film. Only one high quality study addressed this issue.
- in the same study, the use of PACS workstations did reduce the time between image acquisition and the start of treatment for a subgroup of patients (thought to be newly admitted or emergent cases). The study suggested that the clinicians saved time by interpreting more of the images themselves, rather than consulting with radiologists. The accuracy of their interpretations was not studied.
- PACS can reduce problems with missing images. Significantly fewer images are “lost” with a PACS than with conventional film.
- contrary to concerns about possible slowdowns when reading images on computer screens, most studies (5 of 6) reported that images were interpreted on workstations as quickly, or more quickly, as from film.
- user-friendly workstation design and appropriate user training are significant issues in a PACS environment.
- an unintended consequence of “filmless” radiology is that computer maintenance, back-up, and support become mission-critical elements of patient care delivery.

Impact on patient outcomes: no high quality published studies demonstrated that the use of PACS improved patient outcomes. Demonstrating changes in patient outcomes attributable to diagnostic imaging is known to be difficult.

Impact on cost: the recent published literature includes few empirical economic evaluations of PACS. The economic literature consisted largely of discussions of hypothetical cost savings and simulations (models) of potential impacts of PACS on costs.

- most of the empirical research consisted of cost identification studies done from the perspective of the radiology department. Their use of *ad hoc* methodologies limited the usefulness of their findings.
- two published reviews of early economic evaluations of PACS demonstrated that there was wide variation in the reported costs and benefits of PACS, with no consistency in the conclusion as to if, or which, PACS would be less costly or more cost-effective than conventional film imaging. Equipment costs and types of personnel costs included in the studies were the major sources of variance.

- no recent economic evaluations were identified which met the inclusion criteria of this report. Recent empirical studies of PACS costs demonstrated a continued use of *ad hoc* methodologies and simulation, despite the increased availability of evaluable PACS installations, increased maturity of the technology, and increased emphasis on the need to perform high-quality economic evaluations in health care.

- the literature did reflect a growing awareness of the complexity of the evaluation process needed for medical informatics technologies such as PACS. It further suggested that studies using improved methodologies are underway, and should be concluded in the near future. These will hopefully provide useful data to guide policy and research.

Conclusions and Recommendations

Published data do reflect the changing nature of the technology and the complexity of implementing and assessing infrastructure changes. In some cases, published studies also suggest how to ask the appropriate questions and how to design stronger studies. However, the **published evidence does not answer critical questions about the productivity, efficiency or cost-effectiveness of PACS.**

Suggested areas for further planning and research include:

- **clarify and disseminate** strategic goals and objectives for VA telemedicine activities.
- **strengthen the oversight and coordination** of PACS and other telemedicine activities within VA and among federal agencies. The design and implementation of coordinated multi-site assessments of PACS should be included in these efforts.
- further integrate management and economic research methods into PACS assessments to better capture the costs and effects. The literature suggests that we are beginning to understand how to study information technologies, and some useful techniques have been developed in the fields of economics and organizational research. Implicit in this is the need to **strengthen the working relationships among health care economists, management researchers, and clinical researchers within VA.**

- **practice evidence-based decision-making and purchasing**. Require (and fund) more rigorous evaluation of effectiveness, and cost-effectiveness, before broadly disseminating PACS.
- **assess information technologies from an institutional perspective**. Many cost evaluations reflect the traditional focus of information systems and capital expenses on individual departments (usually radiology). In an integrated health care delivery system, the costs and potential benefits associated with PACS, and with the information infrastructure upgrades needed for their implementation, are likely to be system-wide.
- continue to **support test beds for the refinement and evaluation of commercial PACS and in-house image management systems**. While the technology appears to be well-aligned with the mission and strategic objectives of our health care system, credible empirical evidence of efficiency, effectiveness, or economic benefit are still needed.

TABLE OF CONTENTS

I. INTRODUCTION 1
- A. PACS and its significance to VA 1
- B. Significant issues in the clinical use of PACS 2
- C. Challenges in assessing PACS and other medical information technologies 3

II. METHODS 3

III. OVERVIEW OF PACS AND DIGITAL IMAGING 6
- A. Historical perspective 6
- B. Digital imaging background 6
- C. Alternatives to PACS 7
- D. Standards and reimbursement 7

IV. RESULTS OF THE LITERATURE SEARCH 9

V. PUBLISHED FINDINGS AND DISCUSSION 10
- A. Issue 1: Diagnostic accuracy 10
- B. Issue 2: Work process efficiency 16
- C. Issue 3: Effects on clinical care and patient outcomes 22
- D. Issue 4: Cost savings 23

VI. CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS 25

VII. REFERENCES AND SELECTED BIBLIOGRAPHY R-1

VIII. APPENDIX 1: Systematic Search and Appraisal Methods A1-1

IX. APPENDIX 2: Evidence Tables for Diagnostic Accuracy Studies A2-1

X. APPENDIX 3: Evidence Tables for Clinical and Production Efficiency Studies . . A3-1

XI. GLOSSARY G-1

LIST OF TABLES AND FIGURES

Figure 1 Systematic Review Protocol 4
Table 1 Classification of the literature by main topic(s) studied 9
Table 2 *Overview of the literature*: Diagnostic accuracy, film vs. workstation (WS) 12
Table 3 *Overview of the literature:* Diagnostic accuracy, film vs. laser-printed digital image 13
Table 4 *Overview of the literature:* Diagnostic accuracy, laser-printed digital image vs. WS 15
Table 5 *Summary of the literature:* Production efficiency 17
Table A1.1 Evidence-based medicine criteria for studies of diagnostic tests A1- 3
Table A1.2 Methodologic quality rating for diagnostic accuracy studies A1- 3
Table A1.3 Continuum of study designs and their causal implications A1- 4
Table A1.4 Defining and rating of time measurement techniques A1- 5
Table A1.5 Stages of economic evaluation A1- 7
Table A2.1 *Summary of the literature*: Diagnostic accuracy, film vs. workstation (WS) A2- 1
Table A2.2 *Summary of the literature*: Diagnostic accuracy, film vs. printed digital image A2- 8
Table A2.3 *Summary of the literature*: Diagnostic accuracy, soft vs. hard copy A2-12
Table A2.4 Grading the quality of diagnostic accuracy studies A2-15
Table A3.1 *Summary of the literature*: Clinical and production efficiency A3- 1
Table A3.2 *Summary of the literature*: Image interpretation time A3- 6

I. INTRODUCTION

A. PACS and its significance to VA

Picture archiving and communication systems (PACS) are computerized information systems for the management of medical imaging. Although PACS vary widely in equipment configuration and clinical application, they all integrate subsystems designed to acquire, store, distribute and electronically display medical images (Tucker, 1995). Proponents have looked to advanced computer technologies, such as PACS, to increase the efficiency of radiology services, and to improve the quality of care while reducing costs. Several characteristics of radiology services that appear to support this vision, along with some underlying assumptions, are listed below:

- Radiology is resource-intensive, with diagnostic imaging comprising about 5% of the national health expenditure in industrialized countries (Banta and Luce, 1993). (*assumption*: Increased efficiencies could lead to substantial savings.)

- Diagnostic imaging is ordered to provide physicians with information needed to make diagnostic and treatment decisions for selected conditions. (*assumption*: More rapid access to imaging results could lead to quicker diagnosis and clinical action, shorter length of stay, improved patient outcomes, and reduced costs.)

- Even basic radiology data can be generated by computer and stored directly into computers. Nearly a third of the imaging now done in the U.S. is generated by computerized equipment such at CT, MRI, nuclear medicine, and ultrasound scanners. (*assumption*: Radiology is already partly computerized. Full computerization could integrate the service, streamline work flow, and reduce costs.)

Historically, VA has played a major role in the integration and dissemination of information technology into medicine. Consistent with this history, VA decision makers are being asked to fund the purchase of a variety of PACS and related medical information systems. Several types of PACS are already in use within VA, and new PACS are being acquired and installed at many VA Medical Centers.

VA has also developed its own image management system, the VISTA (Veterans Health Information Systems and Technology Architecture) Integrated Imaging System. This system, which is still being refined and evaluated, includes a complete infrastructure for the collection, storage, and retrieval of images, and linkage to patient information within the hospital's integrated information system. The VISTA Integrated Imaging System can be used alone or can be interfaced with commercial PACS to archive images, to integrate them with the evolving electronic patient medical record, and to provide access to images from workstations throughout VA (Dayhoff, 1997). VA and the Department of Defense (DoD) are currently involved in several collaborative telemedicine efforts that include integrating PACS into their systems for patient care delivery. VA has identified an integrated information system as a key component in the transformation of its health care system (Kizer, 1996).

System-wide implementation of PACS, along with the needed infrastructure upgrades, represents a multimillion dollar resource commitment for VA. VA leadership, along with other health care planners worldwide, recognizes the need to evaluate PACS to inform decision making and purchasing (Banta and Luce, 1993; Keen, 1994; Enning, 1994). Assessments of PACS are presently being conducted at the West Los Angeles VAMC, at the Philadelphia VAMC, and at the Baltimore VAMC. As part of VA's evaluation efforts, the

Office of Research & Development requested that the Management Decision and Research Center (MDRC) frame questions for potential future service-directed research.

In respond to this request, the MDRC conducted a systematic review of the literature to define significant issues, and to record the existing knowledge on those issues. The Office of Research & Development coordinated this effort with the Hybrid Open Systems Technology Program (HOST). Based on the systematic review's findings, both programs acknowledged the need for and supported the production of a cost analysis template. The template will be disseminated by the MDRC in a companion document to this report.

B. Significant issues in the clinical use of PACS

This report is based on a review of the published scientific literature. It focuses on studies that provided empirical data, rather than opinion, regarding PACS' impact on diagnostic accuracy, clinical care, and organizational performance. Specifically, the significant issues selected by the MDRC for review include:

1. How does the **diagnostic accuracy** of interpreting digital images on a workstation or printed to film (either or both available with PACS) compare with that of conventional film-screen radiology?

2. How does **work process efficiency** in a PACS environment compare with that of a conventional, largely film-based, radiology department?

3. What are the effects of PACS on clinical care and **patient outcomes**?

4. What **cost savings** accrue from changes in processes and outcomes of care that can be attributed to PACS?

This topic selection reflects both the content of the published literature as well as the MDRC's charge to identify important technology assessment issues in the field. The literature synthesis will summarize the best available empirical evidence in published findings, and will complement the technical research and development efforts already underway within VA.

Underlying the selection of critical issues is Donabedian's model of the relationships between quality, cost, and health. Donabedian (1982) defined two types of efficiency in health care. Clinical efficiency relates to the strategy of care chosen by the physician. The most clinically efficient care produces the best quality of care per dollar spent. Production efficiency refers to processes related to how the services are produced. This includes activities such as reporting of findings to providers, or the use of personnel with the appropriate level of training for a task. Improvements in production efficiency allow for the delivery of the current level of quality at lower cost, or for the production of more care with no change in quality or cost. The largest increments of health (or more quality) per dollar of expenditure are gained by increasing *both* clinical efficiency and production efficiency. This model assumes that better structures and processes result in better outcomes, and that resources freed up by improved structures or processes are used to provide more or better health care (Persson, 1993).

C. Challenges in assessing PACS and other medical information technologies

Clinical applications of PACS are designed to improve both clinical care and organizational processes. While some technical aspects of system performance can be measured directly, it is difficult to design studies that show relationships between the installation of a PACS and measurable changes in organizational process or clinical outcomes. The literature confirms that PACS, like other applications of medical informatics technologies, have been difficult to evaluate. Reasons for this include:

- medical information technologies should be assessed on multiple levels:
  - technical capabilities (assessment of PACS components as a set of hardware, software, and infrastructure tools),
  - as an integrated system performing a clinical task, and
  - as an integral part of the health care delivery system (a tool used to support improvement in the clinical and production activities in partnership with patients, providers, and the organization);
- classical randomized clinical trials, the "gold standard" for evaluation, are rarely feasible for systems such as a PACS, which are very costly to install, require extensive infrastructure development, and impact multiple levels of an organization;
- the diversity in both equipment architectures and clinical implementations makes the technology itself difficult to define;
- PACS and related infrastructure technologies are rapidly evolving; evaluation becomes an iterative process in which early assessment findings are used to guide systems improvements, which must then be reevaluated ("moving target" problem).

II. METHODS

The MDRC performed a *systematic review* of the published literature. A systematic review uses a structured approach that is designed to limit bias and to improve the accuracy of the conclusions that can be drawn from the available published information (Guyatt, 1995).

An overview of the protocol used for this report is presented below. Appendix 1 contains a detailed discussion of the methodology. Technical terms used in the report are defined in the Glossary.

Figure 1. Systematic Review Protocol

A. Define a focused clinical question

B. Use appropriate and explicit inclusion criteria

1. **General criteria for all studies:**
   - English language articles reporting primary data obtained in a clinical setting, or a high quality review of such articles;
   - study design and methods clearly described;
   - study not superseded by subsequent publication with the same purpose from same group;
   - publication date of 1990 or later, to reflect the capabilities of current technology.

2. **Studies of diagnostic accuracy** must also meet most or all of the evidence-based medicine criteria for studies of diagnostic tests (Hayes and Sackett, 1995):
   - clearly identified comparison groups, one or more of which is free of the target disorder;
   - either an objective diagnostic standard or a contemporary clinical diagnostic standard with demonstrably reproducible criteria for subjectively interpreted components;
   - interpretation of the test without knowledge of the diagnostic standard result;
   - interpretation of the diagnostic standard without knowledge of the test result.

3. **Studies of cost** must meet the general inclusion criteria, and must also meet Stage II economic evaluation criteria as defined by the Health Economics Research Group (Sculpher et al., 1995).

   Stage I, Early Developmental: to assess the potential cost-effectiveness of a new technology.
   - begin after making preliminary identification of possible links between process and outcome;
   - use a range of methods, including systematic reviews, *ad hoc* surveys, and early patient data, if available;
   - data sources and study designs likely to provide weak evidence.

   Stage II, Maturing Innovation: to make preliminary estimate of cost-effectiveness.
   - undertaken when patient based data on costs and outcomes is more available;
   - usually uses uncontrolled clinical series, small RCTs, and modeling techniques with sensitivity analyses;
   - data sources and study designs will not provide strong evidence.

   Stage III, Innovation Close to Diffusion: to collect key data for cost-effectiveness analysis.
   - modeling techniques may be used to build framework for the synthesis;
   - may incorporate economic data collection into RCTs;
   - may synthesize effectiveness data from RCTs together with economic data from other sources;
   - data sources will provide strong evidence of effectiveness, but evidence may not be generalizable to other settings.

   Stage IV, Moving into Routine Clinical Practice: to apply Stage III findings to regular clinical practice, and to local settings.
   - may require long-term data collection, and may not be feasible or justified;
   - would provide the strongest evidence of cost-effectiveness in a clinical setting.

C. Conduct a comprehensive literature search

D. Appraise the validity of the individual studies in a reproducible manner

1. **General criteria for all studies:**
   apply inclusion criteria as described above.

2. **Studies of diagnostic accuracy** are then rated using a methodologic quality rating:

   Grade A, high quality evidence: based on studies with broad generalizability with no significant flaws in their research methods;
   Grade B, good quality evidence: based on studies with a narrower spectrum of generalizability and with only a few, well-described, flaws in their research methods;
   Grade C, weak evidence: based on studies with several flaws in their research methods, small sample size, or incomplete reporting;
   Grade D, non-contributory evidence: based on studies with multiple flaws in research methods or reports of opinions unsubstantiated by data.

3. **Studies of processes and outcomes of care** are first evaluated based on their study design, and those providing the strongest evidence of a relationship between the intervention and the outcome of interest are included:

   Firm evidence: based on RCTs or systematic reviews of RCTs;
   Moderately firm evidence: based on prospective cohort studies;
   Highly suggestive evidence: based on historical cohort studies;
   Moderately suggestive evidence: based on case-control studies;
   Suggestive evidence: based on cross-sectional studies;
   Speculative evidence: based on case histories and anecdotes.

4. **Work process studies** that measured time to perform tasks are further rated based on the strength of the technique used to perform the measurements:

   Most accurate: time and motion studies;
   Highly accurate with sufficient number of observations: activity sampling;
   Less accurate than measurements by external observer: self-recording or self-reporting;
   Fairly accurate: administrative data analysis.

5. **Economic evaluations** that meet the general inclusion criteria and the criteria for a Stage II economic evaluation defined by the Health Economics Research Group are then appraised according to the design(s) of the underlying studies from which the data used in the analyses were derived.

III. OVERVIEW OF PACS AND DIGITAL IMAGING

A. Historical Perspective

Picture archiving and communication systems (PACS) are high speed computer network systems for the management of medical imaging. PACS vary in size and configuration, but all provide the following capabilities:

- **storage** of digitized radiology images into a central image database (archiving);
- **access** to these images from computer workstations;
- **rapid transmission** of images from the radiology acquisition devices to the archive and from the archive to the workstation (Bakker, 1988).

First introduced in 1982, PACS were made feasible by advances in two fields: medical imaging and information technology. Since the invention of computed-emission tomography (CT) in the early 1970's, there has been an increasing reliance on computers, rather than x-ray film, to produce radiological images. Recent advances in microcomputer technology and telecommunication capabilities have created the opportunity to evolve new systems for storing and using medical images, and for integrating them into administrative and clinical information systems.

A fully integrated information system is considered a key infrastructure requirement in an efficient and effective health care system. PACS can support the integration of imaging into the information infrastructure on many levels. A "mini-PACS" can be used within one department, usually radiology, an ER, or an ICU. Larger PACS can support the wide distribution of radiology images, and can integrate imaging into a Hospital Information System (HIS) or into computer-based patient records. PACS can be used over networks to connect patients, providers, and facilities over great distances. This last application, called teleradiology (radiology at a distance), is being explored as a way to increase access to care.

The size and configuration of each PACS installation can be customized because PACS are actually comprised of equipment subsystems networked together (Becker, 1994). Equipment choices, the structure of the connecting network, and the level of integration of imaging into other information systems can all vary. The heterogeneity of PACS equipment and the variability in the level of integration of PACS into the organization's electronic information system have contributed to the difficulty in evaluating PACS.

B. Digital imaging background

A brief overview of digital imaging is included to help clarify some issues discussed in the literature review.

Historically, radiologic images were made by radiating photographic film. This produced high contrast images which were viewed on a backlit view box, often located in a dedicated "reading room." These conventional film images are fundamentally different from the digital images produced by computer systems. Conventional x-ray film images can display any value of light intensity and can show any amount of separation between objects. They contain information that can vary continuously. This is known as analog information. Computers use digital information. They store and process information in discrete numeric

units, that is, as discontinuous pieces of information. For computers to process image information, analog (continuous) signals must be converted to digital information. The design of the digital image capture and display systems help to determine just how closely the digital data can represent the original analog signal.

The more information (bits) associated with each picture element (pixel), the more closely the computer image can approximate an analog image. One bit associated with one pixel can represent two values, black and white. Two bits can distinguish four values: black, white, and 2 shades of gray, 8 bits = 256 gray levels, and so on. If color is used in the image, more bits per pixel are needed.

Two parameters commonly used for comparing image quality are spatial resolution and contrast resolution. Spatial resolution, the ability to show physical separation between objects, is determined by the number of pixels per square inch. Contrast resolution, the ability to distinguish small differences in intensity, is determined by the number of bits per pixel.

Digital images can be displayed on a workstation for viewing as "soft copy." Imaging workstations may be categorized based on display capabilities: low-resolution (512 x 512 pixels); medium-resolution (approximately 1000 x 1000-1600 pixels); and high-resolution (approximately 2000 x 2000-2500 pixels or greater). Workstations also vary in the capabilities they offer users to adjust the displayed image (such as the ability to magnify, or to change the contrast or brightness settings). Digital images can also be laser-printed to film. This "hard copy" of the digital image resembles a conventional x-ray film, and is read on a light box like a conventional x-ray film.

C. Alternatives to PACS

Clinical evaluations of PACS are typically based on comparisons to the traditional alternative technology, a conventional radiology system with analog film being viewed on a light box and stored in a file room.

Some studies include comparisons among PACS workstation imaging, conventional film, and digital images laser-printed to film (or between any two of these three). Although PACS could be used with laser-printing equipment and all imaging could be printed to film, many of the hypothetical benefits of PACS are based on the assumption that images would be viewed on workstations and stored electronically, creating a truly "film-less" radiology service.

D. Standards and reimbursement

High performance computing and communications technologies, including PACS, have the potential to increase access to health care, reduce costs, and improve efficiency. When implementing such technologies, health care delivery systems are confronted with the need to address the implications of these technologies regarding compliance with privacy laws, licensing and credentialing regulations, data ownership, malpractice liability, and reimbursement processes (Bradham, Morgan, and Dailey, 1995). Many of these issues are being actively addressed by legislative and regulatory bodies and also by professional and provider organizations. A discussion of the societal implications of telemedicine and related technologies is presented in the Telemedicine Report to Congress (U.S. Department of Commerce, 1997).

Provider and equipment standards for remote radiology were developed by the American College of Radiology (ACR, 1994). They recommended that the physician providing an official image interpretation should have documented training in diagnostic radiology and an understanding of the digital imaging technology, equipment, and processes. They further recommended that physicians maintain licensure at both the transmitting and receiving sites when interpreting images remotely. Equipment recommendations included the use of a 2K x 2K x 12 bits array or better for digitizing an image, and a 2K x 2K x 8 bits array or better for workstation displays to be used to produce the official reading of digitized radiographic films and computed radiography. A 0.5K x 0.48K x 8 bits array or better was recommended for small-matrix systems such at CT and MRI.

In 1995, the Health Care Financing Administration (HCFA) approved physician reimbursement for the interpretation of electronically transmitted radiological images, if the interpretation contributed to patient diagnosis or treatment and if the physician was licensed to perform the service in the state in which it is furnished (Federal Register, December 8, 1995). HCFA has left it up to the individual insurance carriers to decide whether to reimburse for teleradiology and similar services not requiring face-to-face consultation with the patient (HCFA, 1997).

IV. RESULTS OF THE LITERATURE SEARCH AND APPLICATION OF INCLUSION CRITERIA

Six hundred thirty five unique citations and abstracts were screened by the authors of this report. Copies of full articles were obtained for those citations judged to be potentially relevant. Twenty two articles (3.5% of those identified in the search) met the inclusion criteria for this report. The eligible articles fell into two broad categories: assessing PACS as a diagnostic tool used in a clinical setting (assessment of diagnostic accuracy), and assessing PACS as an infrastructure element (assessing the impact on production efficiency, clinical performance, or systems costs). Overall, the included studies yielded the following numbers of observations about these issues:

Table 1. Classification of the literature by main topic(s) studied*

Research Issues	Comparison	# of Observations
Assessment of PACS as a diagnostic tool used in a clinical setting		
Diagnostic accuracy	• conventional film vs. digital image viewed on workstation	11
	• conventional film vs. digital image printed to film	6
	• digital image printed to film vs. digital image viewed on workstation	4
Assessment of PACS as an infrastructure component		
Process efficiency	• time to generate and deliver images	2
	• time to retrieve recent images	4
	• time to retrieve older images	2
	• time to first encounter with imaging information	1
	• time to interpret images	6
	• ability to access images (PACS down time)	1
	• percent of images lost	1
Patient care / outcomes	• time to initiate clinical action	1
	• effect on outcomes of care	0
Cost [savings]	• high-quality reviews of early economic evaluations	2
	• high-quality Stage II economic evaluations	0
Total number of observations		41 (in 22 eligible studies)

*Note that studies addressing multiple issues are counted in all categories studied

Diagnostic accuracy studies comprise most of the literature reporting primary empirical data on the clinical application of PACS (Table 1). This is an expected finding. Research is first done to learn if a technology is accurate enough to justify its use in a clinical setting, then studies of the clinical and organizational impact (including cost) are undertaken.

The distribution the types of studies in the PACS literature is consistent with a recent analysis of telemedicine literature. Tohme (1996) reported that a bibliographic search revealed that most of the published literature on telemedicine consisted of discussions of the potential of the technology and methods or planning pieces. Empirical tests of telemedicine in a clinical setting, using *any* study design, accounted for less than 30% of the literature. Four percent of published studies evaluated organizational issues, and only 2% studied costs.

V. PUBLISHED FINDINGS AND DISCUSSION

The following section of the report is organized by topic. The discussion of each significant issue is accompanied by tables containing an overview of the evidence provided by studies that met the inclusion criteria. If a study addressed more than one of the questions relevant to this report, it was included in more than one table. Where published quality standards were identified, studies were graded for methodologic quality. Full data collection tables, from which the overview tables were derived, are presented in the appendices along with the methodologic quality grading tables.

A. ISSUE 1: How does the diagnostic accuracy of digital imaging compare with that of conventional film-screen imaging?

1. Published findings

From the articles retrieved, 15 reports were identified which met the inclusion criteria for studies of the diagnostic accuracy of digital imaging. Overview tables summarize the diagnostic studies that compared:

- conventional film to digital images viewed as soft copy on workstations (Table 2);
- conventional film to digital images viewed as hard copy on film (Table 3); and
- hard copy to soft copy of digital images (Table 4).

Appendix 2 contains the full data collection tables from which Tables 2 through 4 were derived. The appendix also contains a table classifying the quality of these 15 studies (Table A2.4).

None of the identified studies were classified as Grade A, that is, a study with broad generalizability to a variety of patients and no significant flaws in the research methodology. The most common methodologic problems identified were: limited spectrum of patients (typically over-representation of selected disease states); retrospective study design; and insufficient size to detect differences in accuracy if they truly existed.

Most of the studies of diagnostic accuracy used receiver operating characteristic (ROC) curves to analyze performance. This is a particularly useful analytic method, because it explicitly acknowledges the role of the observer in the process of image interpretation. Any imaging technology used for making a diagnosis requires (at present) that a human observer interprets the image. The ability to interpret an image varies over time for individual readers, and varies among readers.

***Conventional film versus digital images viewed on a workstation*:** Table 2 summarizes the 11 studies that compared conventional film to digital images viewed as soft copy on a workstation. Most of the studies published from 1990 to the present assessed moderate or high-resolution imaging.

One recent study assessed the diagnostic performance of low-resolution workstations for use with chest and bone images. The study findings did not support the use of low-resolution monitors for making primary diagnoses (Paakala, 1991).

The results from studies using moderate resolution equipment were mixed, and no consistent relationship between methodologic quality and the study findings was noted. Ackerman (1993) and Goldberg (1993) both reported that the overall accuracy of interpretations was significantly higher for conventional film than for workstation interpretation. Scott (1995) reported that, while overall accuracy was significantly higher for conventional film than for workstation images, these differences varied with the diagnostic difficulty of the disorder and with the training of the reader. Korsoff (1995) found that moderate-resolution workstations were equivalent to convention film for the primary diagnosis of selected pulmonary diseases, but that the workstations studied were not adequate for the primary diagnosis of selected subtle diseases. Using somewhat higher-resolution equipment, studies by Elam (1992) and DeCorato (1995) reported equivalent accuracy for conventional film and workstation interpretation of a variety of image types. However, Slasky (1990) and Thaete (1994), in the two studies graded to have only a few methodologic flaws (Grade B), reported that conventional film was either equivalent to, or more accurate than, workstation imaging. Their relative accuracy depended on the nature of the diagnosis being made.

Findings with high-resolution workstations (approximately 2K x 2K pixel matrices) also varied. Using pediatric images, Franken (1992) reported “no appreciable difference” in accuracy of interpretation between conventional film and images viewed on a high-resolution workstation. Using adult chest images, which contain more information than the small pediatric images, Cox (1990) found that the accuracy of conventional film was either better, worse, or equivalent to that of workstation images, depending on the disorder being diagnosed.

***Conventional film versus digital images printed to film*:** Table 3 summarizes the 6 studies that compared conventional film and digital images printed to film. Studies assessed the performance of systems using approximately 2K x 2K or better digital image resolution.

Findings were mixed. Conventional film outperformed full-sized digital film, but was equivalent to 2/3 sized digital film, in the study by Kondoh (1994) which contained multiple methodologic flaws (Grade D). Slasky (1990) and Thaete (1994), the two most rigorously designed studies identified (Grade B), reported that relative accuracy of convention versus digital film varied with the condition being diagnosed. Slasky (1990) reported that conventional film was more accurate than digital film for the diagnosis of pneumothorax, but equivalent for the other conditions studied. Using higher resolution imaging, Thaete (1994) reported that conventional film was more accurate than digital film for the detection of interstitial disease, but equivalent for the detection of pneumothorax and other conditions assessed. Elam (1991) reported no significant differences in accuracy, sensitivity, or specificity of detection of pneumothorax between conventional film and digital film printed in either small or large format. Cox (1990) reported no significant difference between conventional film and digital film for the detection of 8 of the 9 pulmonary conditions studied, but noted that digital film was more accurate than conventional film for detecting parenchymal masses. Yoshino (1992) reported that diagnostic performance varied with reader experience. The more experienced readers, but not the less experienced

Table 2: *Overview of the literature*
Diagnostic accuracy studies: comparison of conventional film to digital images viewed on a soft copy (on a workstation)

Notes:

- *Table A2.1 contains a full summary of the studies in this table, and their methodologic quality grading is reviewed in Table A2.4 (both in Appendix 2).*
- *All studies are observational, and most used a case-control design with randomized presentation of images to readers. Studies by Paakkala, Goldberg, and DeCorato used a weaker experimental design, that of a case series in which patients determined to be free of disease were in the case series, and served as internal controls.*
- *ROC (receiver operating characteristic) curves were used to compare diagnostic accuracy except in studies by Paakkala and DeCorato.*
- *The diagnostic standard (gold standard) for studies included in this table was the confirmed diagnosis by biopsy, CT, expert opinion based on clinical evidence, or comparison to conventional analog film.*
- *Since blinding of the readers was not specifically indicated in some studies included in the table, those studies do not fully meet evidence-based criteria for diagnostic test evaluation. Because of study design limitations, evidence presented in this table can be considered suggestive or moderately suggestive.*
- *Studies by Slasky, Cox, Elam, and Thaete also included comparisons to digital imaging printed to hard copy (film). These findings are reported in subsequent tables.*
- *Ackerman, 1993 and Scott, 1995 were performed at the same institution, but address different diagnostic problems.*

Study	Grade	Image Type	Digitized image resolution (if indicated)	Workstation resolution	Film > WS	Film = WS	Varies with diagnosis	Varies with reader experience
Paakkala	**C**	chest and bone		512 x 512 x 8	+			
Ackerman	**C**	adult/pediatric chest and bone		1280 x 1024	+			
Korsoff	**C**	chest	1024 x 1024 x 12 (moderate resolution)	1280 x 1024 x 8			+	
			2048 x 2048 x 12 (high resolution)	1280 x 1024 x 8			+	
Elam	**C**	adult chest, pneumothorax	1760 x 2140	1024 x 1536		+		
DeCorato	**D**	any ER image: 693 film, 118 CT, 1 MRI	2040 x 2056 x 12	1200 x 1600		+ (for 95% of cases)		
Scott	**C**	chest, musculoskeletal, abdominal		1200 x 1600			+	+
Slasky	**B**	adult chest	2048 x 2400 x 12	1536 x 2048 x 8			+	
Thaete	**B/C**	adult chest	4096 x 5000 x 12	1536 x 2048 x 8			+	
Franken	**C**	newborn chest	1024 x 1024	2000 x 2000		+		
Goldberg	**D**	77% adult/34% pediatric chest, bone, urinary tract	1684 x 2048 x 12	2048 x 2560 x 8	+			
Cox	**C**	adult chest	2048 x 2048 x 12	2560 x 2048			+	

Reported outcomes as listed in the column headings:

Film > WS: accuracy of detection of a disorder is significantly higher for conventional film than for soft copy of digital image displayed on a workstation.

Film = WS: accuracy of detection of a disorder is comparable for conventional film and for digital image displayed on a workstation.

Varies with diagnosis: accuracy of detection of specific subtle disorders is significantly higher for conventional film than for workstation; accuracy of detection of other abnormalities comparable for both modalities. In the study by Cox, accuracy of detection of one abnormality was significantly lower for conventional film than for digital image displayed on workstation.

Varies with reader experience: senior radiologists performed better with conventional film than with digital image displayed on a workstation for some diseases, but performed equally well in both mediums for the diagnosis of other diseases. Less experienced readers performed comparably using either display mode.

Table 3: *Summary of the literature*
Diagnostic accuracy studies: comparison of conventional film and digital images viewed as hard copy (film)

Notes:

- *Table A2.2 contains a full summary of the studies in this table, and their methodologic quality grading is reviewed in Table A2.4 (both in Appendix 2).*
- *All studies used a case-control design and randomized presentation of images to interpreters.*
- *ROC (receiver operating characteristic) curves were used to compare diagnostic accuracy.*
- *The diagnostic standard (gold standard) used was confirmed diagnosis by biopsy, CT, or expert opinion based on review of images and/ or medical records.*
- *Studies included in this table met most or all of the evidenced-based criteria for diagnostic test evaluation; studies by Elam, Yoshino, and Kondoh do not explicitly indicate that readers were blinded. Because of study design limitations, evidence presented in this table can be considered suggestive or moderately suggestive.*
- *Studies by Slasky and Thaete were performed by different departments in the same institution, using different subjects and methods.*

Study	Grade	Image Type	Digital Image Resolution	Digital film size	CF > DF	CF = DF	Varies with diagnosis	Varies with reader experience
Elam	C	adult chest: pneumothorax	1760 x 2140	small (7” x 8.5”)		+		
				full-sized (14” x 17”)		+		
Kondoh	D	adult chest	2000 x 2000	2/3 -sized		+		
				full-sized	+			
Cox	C	adult chest	2048 x 2048	full-sized			+	
Yoshino	C	cervical spine	2048 x 2048	56%-sized				+
Slasky	B	adult chest	2048 x 2400	reduced 4%			+	
Thaete	B/C	adult chest	4096 x 5000	reduced 9.3%			+	

Reported outcomes as listed in the column headings:

CF > DF: accuracy of detection of a disorder significantly higher for conventional film than for digital image laser-printed to film.

CF = DF: accuracy of detection of a disorder comparable for conventional film and for hard copy of digital image (film).

Varies with diagnosis: accuracy of detection of specific subtle disorders is significantly higher for conventional film than for digital image hard copy (film); accuracy of detection of other abnormalities is comparable for both modalities. In the study by Cox, accuracy of detection of one abnormality was significantly lower for conventional film than for digital image laser-printed to film.

Varies with reader experience: senior radiologists performed better with conventional film than with digital image displayed on a workstation for some diseases, but performed equally well in both mediums for the diagnosis of other diseases. Less experienced readers performed comparably using either display mode.

radiologists, were significantly more accurate when using conventional film than workstation images.

Digital images printed to film versus digital images viewed on workstations: Table 4 summarizes the 4 studies in which the diagnostic accuracy of digital images viewed as hard copy was explicitly compared to digital imaging viewed as soft copy. All studies contained several flaws in methods (Grade C). The quality of imaging and film size varied across studies.

Three studies reported equivalent diagnostic accuracy when using either hard or soft copy of digital images for viewing abdominal CT images, adult chest (pneumothorax), or pediatric chest images. Cox (1990) reported that digital film was as accurate as workstation viewing for the detection of five of the pulmonary conditions assessed, but that digital film was more accurate than workstation viewing for detecting three pulmonary conditions.

2. Discussion

Published studies of the diagnostic accuracy of digital imaging relative to conventional film-screen imaging did not strictly follow widely known principles of study design for avoiding bias in the evaluation of a diagnostic test. They used study designs, methods, or analyses that would be expected to overestimate the accuracy of digital imaging and to reduce the validity and generalizability of the findings. Equipment and methods varied significantly, making meta-analysis inappropriate and synthesis challenging.

The use of low resolution displays (approximately .5K x .5K) for making primary diagnosis from chest and bone images was not supported by the evidence provided in one identified study. This is consistent with the early literature in digital imaging, with identified concepts in visual perception (Kundel, 1986), and with the professional standards of the American College of Radiology (1996).

Findings from studies of moderate and high resolution imaging were mixed, and, because of methodologic and reporting limitations of the studies, it was difficult to identify the main contributors to the variability in findings. The most methodologically sound studies (Slasky, 1990 and Thaete, 1994) both reported that conventional film was either equivalent to, or more accurate than digital images viewed either on a workstation or printed to film, depending on the nature of the diagnosis being made. Most of the available data suggest that the diagnostic accuracy of digital imaging continues to improve. However, available data suggest that conventional analog imaging is more accurate than digital imaging, particularly for making a primary diagnosis of subtle manifestations of diseases that require high-resolution imaging.

Digital images are frequently printed to film for viewing and archiving, even when workstations are available for image viewing. This restores the “look and feel” of conventional analog film, but eliminates the ability to use workstation tools to improve the displayed image. While printing digital images to film does cause the loss of some of the organizational efficiencies thought to be inherent in PACS, data suggest that printing does not degrade the diagnostic accuracy of the images. One study (Cox, 1990) reported that digital film was more accurate than workstation imaging for three of the pulmonary conditions assessed. This finding may be the result of human or technical factors.

Prior experience in reading images on workstations and willingness to use workstations interactively to improve the quality of the imaging were factors discussed, but not quantified, in the literature. Non-uniform user workstation experience and limited use of workstation tools to enhance the displayed images may have reduced relative performance on

Table 4: *Overview of the literature*
Diagnostic accuracy studies: comparison of digital images viewed as hard copy (film) to digital images viewed as soft copy (on a workstation)

Notes:

- *Table A2.3 contains a full summary of the studies in this table, and their methodologic quality grading is reviewed in Table A2.4 (both in Appendix 1).*
- *The studies included in this table used a case-control design and randomized presentation of images to the readers.*
- *ROC (receiver operating characteristic) curves were used to compare diagnostic accuracy.*
- *The diagnostic standard (gold standard) used was consensus of experts based on all available clinical and imaging information, or (in Straub) surgical confirmation when available.*
- *Since blinding of readers is not specifically addressed in the methods sections of three papers summarized in the table, it is unclear if these studies fully met evidence-based criteria for diagnostic test evaluation.*
- *Because of study design limitations, evidence presented in this table can only be considered suggestive or moderately suggestive.*

Study	Grade	Image Type	Digitized image resolution	Digital film size	Workstation resolution (image size)	Hard = Soft	Varies with diagnosis
Straub	C	abdominal CT	512 x 512 x 12	full-sized (14" x 17")	1536 x 2048 x 8; 12 images, 3/4 size on 2 screens	+	
					1536 x 2048 x 8; 12 images, 3/4 size, presented sequentially	+	
					1536 x 2048 x 8 ; 12 images displayed sequentially at 1.4 x size	+	
Elam	C	adult chest, pneumothorax	1760 x 2140	small (7" x 8.5")	1024 x 1536	+	
				full-sized	1024 x 1536	+	
Razavi	C	pediatric chest	2048 x 2048 x 8	157: 8" x 10" 44: 10"x 12" 38: 14" x 17" (analyzed together)	2048 x 2560	+	
Cox	C	adult chest	2048 x 2048 (after compression)	full	2048 x 2560 x 12		+

Reported outcomes as listed in the column headings:

Hard = soft: accuracy of detection of a disorder comparable for digital image displayed as hard copy (film) and digital image displayed on a workstation.

Varies with diagnosis: accuracy of detection using hard copy either better, worse or comparable to accuracy of detection using soft copy, depending on the disorder being viewed.

workstations and contributed to inconsistencies in findings of diagnostic accuracy studies. The role of the reader in the process of image interpretation warrants further quantitative research.

Published studies provide insufficient information to quantify the *clinical* (as opposed to *statistical*) significance of differences in diagnostic accuracy between digital and analog imaging. Some data suggest that the inter-reader and intra-reader variability in accuracy that already exists in the analog environment is larger than some of the differences in accuracy reported between analog and digital imaging. Statistical techniques have been developed for comparing ROC curves that include the assessment of inter-reader and intra-reader variations (Kundel et al., 1996). Applying such methods in a well-designed study could provide valuable information about whether the quality of the performance of diagnostic tasks is changed when PACS are introduced.

Based on the available studies and the application of quality criteria, there has not yet been a definitive demonstration that digital imaging, viewed either on a workstation or printed to film, is equivalent to analog film for making a primary diagnosis of all of the clinical conditions that present in a varied patient population. Data do reflect the recent improvements in digital imaging technology, and some data do suggest that both 2K x 2K hard copy (film) and soft copy (workstation display) are equivalent to conventional film for accurate primary diagnosis of many of the disorders routinely encountered in patient care. The requirements for resolution of image capture and image display systems depend, in part, on the disorder being studied, and larger matrices (smaller pixel size) are needed to visualize the fine detail present in selected abnormalities. Well-designed studies are still needed to show that newer PACS imaging technologies are equivalent to analog film for use in making primary diagnoses in a clinical setting.

B. ISSUE 2: How do work process efficiencies with PACS compare with those in a conventional radiology department?

1. Published findings

From the articles retrieved, 9 studies were identified which presented empirical evidence related to work process efficiency, and which met the inclusion criteria for this report. Four of these studies were primarily designed to assess diagnostic accuracy, and were reported in earlier tables. Table 5 summarizes the strongest evidence presently available for the assessment of work process efficiencies in a PACS environment relative to those in a film-based environment. Appendix 3 contains the full data collection tables from which Table 5 was derived (Tables A3.1 and A3.2).

No relevant randomized clinical trials (RCT) or high-quality meta-analyses were identified. Studies included in the report were case reports, case series, cross-sectional, or case-control studies, all of which provide a weaker level of evidence of the relationship between the intervention (PACS) and the outcomes of interest (change in process) than that would have been provided by a RCT (Table A1.3). Because studies varied widely in design, methodology, and analysis, a meta-analysis of the identified studies was not undertaken.

Seven of the 9 studies used time measurement techniques considered to yield highly accurate data (Warburton, 1992). The most methodologically rigorous data collection was done by Kundel (1996), who gathered time element data using time and motion studies (by videotape and direct observation), as well as computer records and self-reporting to trained interviewers. Gay (1997) also did time and motion studies using direct observation, but lacked confirmatory videotape recordings of activities studied.

Table 5: *Summary of the literature*
Production efficiency: work process for film imaging compared with that of PACS

Notes:

- *Studies included in this table (described in detail in Tables A3.1 and A3.2, Appendix 3) represent the strongest evidence presently available for the assessment of process efficiencies in a PACS environment relative to a film-based environment.*
- *All were observational studies, but designs varied.*
- *Studies which reported image interpretation times were primarily designed to assess the diagnostic accuracy of image interpretation, with the exception of Kato (1995). Lou and Huang (1992) reported preliminary impressions of comparability of diagnostic accuracy between PACS workstations (WS) and film. All other studies assumed comparability of diagnostic accuracy between PACS workstations and film.*
- *The studies varied in systems architecture and methodology. While all included clear descriptions of the elements of the clinical process timed in the study, the process elements measured varied widely, as did the time measurement techniques used. This heterogeneity makes comparisons across studies difficult. While an overview of the studies does suggest process and outcome findings which are valuable for future assessments of production efficiency, the methodological limitations of the studies should be considered when interpreting findings.*
- *When possible, time measurement technique were classified according to the methods described in Table A1.4. Computer documentation of elapsed time was considered to be a highly accurate measure.*

Study	Time Measurement Technique	Comparisons of time spans to perform identified task							Percent of images lost	Radiologist was first source of information
		Generate and deliver images	Retrieve recent images	Retrieve archived images	Image information first accessed	Interpret images	Initiate clinical action	System reliability		
Horii et al., 1992	self-recorded		film = PACS	film > PACS					film > PACS	
Lou & Huang, 1992	self-recorded or self-reported	film > PACS	film > PACS					archived PACS images available 95% of time; images in workstation available 99.97% of time (film not assessed)		
Kundel et al., 1996	multiple techniques, including time and motion study with video	film > PACS			film = PACS		film > PACS			film > PACS
Franken et al., 1992	time and motion study		film = PACS			film < PACS				
Straub et al., 1991	computer record of elapsed time					film = PACS				
Razavi et al., 1992	computer record of elapsed time					film = PACS				
Thaete et al., 1994	computer record of elapsed time					film = PACS				
Kato et al., 1995	time and motion					film = PACS				
Gay et al., 1997	time and motion			film > PACS		film > PACS				

Identifying key work processes and inefficiencies in conventional radiology departments: Scant empirical evidence has been published about the work flow and clinical processes for patient imaging.

In a study of radiology work flow in a conventional film environment, Heymann and Culling (1996) identified 54 distinct activities and 18 incidents of work process passing from one person to another, including the time in which a routine chest x-ray film was ordered on an inpatient to the time the film was re-filed after having been reviewed and reported. Eighty-two percent of the staff time was used in work processes that did not involve technical or medical functioning. No studies were identified which mapped the entire work process in a PACS environment.

One study provided empirical evidence identifying distinct work activities that were major sources of inefficiency in a conventional radiology department. Gay (1997), in a small workflow analysis of CT scans, reported that the retrieval of film folders by the film room personnel acted as a bottleneck for film reading of CT studies. This is consistent with earlier observations about film retrieval problems in conventional radiology departments. Horii (1992) reported that, on their radiology service before the PACS installation, physicians, rather than clerks, often retrieved and personally returned radiology films. Lou and Huang (1992) commented on other bottlenecks in conventional radiology services. They asserted that delays in film retrieval and missing films were two common inefficiencies in most radiology departments, and that these were often caused by competition for films among the clinicians involved in the same patient's care.

Identified work process activities in a PACS environment relative to a conventional film-based environment: Although no clear, comprehensive mapping of radiology work processes before and after PACS installations has been reported, a few studies have attempted to identify key process elements, and to document changes in time taken to perform these tasks.

In the discussion section of his preliminary report of radiology work flow in a PACS versus a film environment, Kato (1995) identified five factors that affect radiology throughput:

1. time taken to transfer physician order information;
2. time to perform the examination;
3. time to transfer image from the radiographic imaging equipment to the radiologist;
4. time required to interpret the image, including the retrieval of previous images for comparison and the selection of settings on the PACS workstation;
5. time to send the reports and images to the referring physicians.

His opinion was that the first two items would remain essentially the same in either a film or PACS environment, but that PACS could reduce the image transfer time and also the time to send images (but not necessarily to send official reports) to referring physicians. Factor 4, image interpretation time, is dependent on both user skill and equipment design, and is the most commonly studied element of change in work flow with the introduction of PACS.

Studies summarized in Table 5 measured six distinct activities and three characteristics of the work process, including factors 3, 4, and 5 identified by Kato (1995). The identified activities and work processes are discussed below.

Lou and Huang (1992) and Kundel (1996) compared several discrete process activities in a PACS environment to those in a film-based environment. Both reported that the time interval

between the completion of an examination and when the image was available for viewing, along with relevant previous examinations, was longer for film images than for workstation images. Lou and Huang did not measure the retrieval time for old films that had been archived off-site.

Findings related to the time taken to retrieve recent images varied. Lou and Huang (1992) reported that the retrieval of recent images was done more efficiently by a PACS workstation than in a film-based system. Horii (1992) and Franken (1992) reported that this interval was comparable in both film and workstation environments. Differences in study design may explain the non-comparability of findings. Both Horii's (1992) and Franken's (1992) studies could have been expected to underestimate the typical time taken to retrieve a recent film image in clinical practice. Horii's case selection included 30 ultrasound (US) images, for which PACS image retrieval was a non-significant 37 seconds faster than film retrieval, and 10 CT images, for which PACS image retrieval was 10 minutes, 2 seconds faster than film retrieval. The overall study findings largely reflected the ultrasound data, whereas the author noted that, in clinical practice, CT films may be requested and signed out more frequently than ultrasound films (Horii, 1992). In the study by Franken (1992), film image access time reflected time taken to access film already mounted on an alternator, but failed to include the much longer time interval taken to retrieve that film from the files in order to mount it.

Gay (1997) and Horii (1992) measured the time taken to retrieve archived images. Horii reported that PACS significantly reduced this retrieval time. Gay reported that, on average, archival retrieval time was 50 minutes shorter for PACS than for film. Gay further noted that the retrieval of film folders from the file room was the rate limiting process (bottleneck) in the workflow for film reading of CT scans.

Six studies summarized in Table 5 assessed the time taken to interpret film versus workstation images. Four of the six studies, using either time and motion studies or computer-documented reading times, reported no significant difference between the time taken to interpret film versus workstation images. One time and motion study reported that, while the interpretation of a newborn chest image took well under a minute regardless of whether film or workstation imaging was used, the interpretation of film was significantly faster than the interpretation of images on PACS workstations (Franken, 1992). A more recent time and motion study (Gay, 1997) reported that the average time for reading a whole body CT scan was 7.108 minutes for film as compared with only 4.205 minutes for PACS workstation reading. When assessed, the diagnostic accuracy of workstation image interpretation was equal to that of film imaging for some, or all, diagnoses studied.

Lou and Huang (1992) asserted that missing films were a common inefficiency in most radiology departments. Horii's study (1992) measured the percentage of cases in which images were lost. Lost images were twice as common for film as for images archived electronically in a PACS (10% versus 5% for recent images). There was actually no true data loss with the PACS; some images had been electronically misfiled because of errors by the technicians.

The introduction of electronic data storage potentially introduces a new type of image retrieval problem, that of computer malfunction. The Lou and Huang study (1992) assessed down time of their PACS to help quantify this potential image retrievability problem. They determined that there was a 95.22% probability that the necessary computer systems would be available to view archived images, and a 99.97% probability that a local workstation would be fully functional and able to retrieve a locally-stored image. No comparable data were collected for a film-based system.

Clinical information acquisition processes in a PACS environment relative to a conventional imaging environment: If imaging information is needed for clinical decision-making, the acquisition of that information is a key step in the clinical work process. Only one identified study, Kundel (1996), assessed the clinical work processes related to the acquisition and use of image information. Kundel (1996) noted that diagnostic information about an image could be obtained in several ways, including direct viewing of the image, "asking a radiologist," or reading a preliminary or final radiology report. His study assessed time to encounter diagnostic imaging information *by any means*, and reported that PACS led to no statistically significant decrease in the time interval between when the imaging exam was completed to the time at which the ordering physician first encountered imaging information. Subgroup analysis showed that, for acutely ill or newly admitted ICU patients, PACS did reduce the time to access imaging information from an average of 4.4 hours for film to 2.5 hours for PACS. This was observed in an environment in which the images were available for viewing in about 10 minutes with PACS and about 1 hour for film.

Kundel also noted a shift in the way imaging information was obtained. He reported a significant decrease in input from radiologists when the workstations were available than in the pre-PACS film-only environment. Most of this was due to a significant decrease in the number of direct contacts with radiologists for imaging information. The decrease in consultations with the radiology staff when workstations were used is consistent with earlier findings reported by the same group (DeSimone, 1988). This produced a situation in which the medical ICU staff used the workstations for primary diagnosis and clinical decision-making. Official radiology interpretations were made later, based on film readings. The impact of the decreased communication on radiologists was not assessed in this study.

2. Discussion

One of the most frequently cited potential benefits of PACS is that these computerized systems will support major increases in efficiency within organizations. The MDRC search of the published literature could identify only limited empirical evidence that addressed the definition of radiology work processes, or which demonstrated changes in those processes related to the use of PACS.

Identified published studies that met the MDRC inclusion criteria used study designs ranging from case reports to case-control studies, all of which yield weaker evidence than randomized controlled trials (Table A1.3). Time measurements were generally collected using methods that would be expected to yield very accurate data (Table A1.4).

Image generation, processing, and retrieval: Data suggest that PACS are more efficient at generating and delivering images than film-based systems. Retrieving recent images was reported to be as fast, or faster, with PACS than in a film-based environment. The retrieval of archived images was significantly faster from a PACS than from film storage. It is difficult to assess the clinical significance of this time savings, since, according to some authors, archived images are infrequently needed for the clinical care of a patient.

Images could be found significantly more often in a PACS than in a film-based environment. The percentage of "lost" cases was reduced from 10% (with film) to 5% (with PACS) for images generated in the prior 6 months. PACS images were not actually lost; they were captured incorrectly and could not be retrieved using patient identifiers. It may be possible to reduce this type of error with appropriate training, but this was not addressed in the study. Film image "loss" reflects, in part, unsuccessful attempts by multiple providers caring for a patient to review the same films. One provider may be reviewing the film without having signed it out. The capacity for multiple clinicians to access the same images at the same time is a major potential benefit of PACS, but one that remains largely unquantified.

There is an unintended consequence of "filmless" radiology that may have significant impact on organizational functioning. PACS images can only be retrieved if the computer system is functioning. All image-related patient care becomes dependent on the equipment and communications systems. In a PACS environment, computer system maintenance, support and training become mission-critical.

Image interpretation: Once images are generated, processed, and retrieved for viewing, the images must be interpreted. The time taken to read an image reflects elements of both equipment functioning and clinician behavior. There had been concern that using workstations to interpret images would be more time-consuming than the familiar process of reading conventional analog film. Six studies were identified which compared the time taken to interpret images from film versus from PACS workstations. Four of the six studies reported no significant difference in image interpretation time between film and workstations. The one study (Franken, 1992) which did report a statistically significant increase in interpretation time when using PACS noted that the workstation reading times, although increased, were still well under a minute. Another study (Gay, 1997) reported that PACS workstation reading times for CT scans were nearly half those of film reading. The lack of agreement between this study's findings and those of other studies might be related to the small study size or the relative greater experience of their readers with PACS workstations.

The use of image processing tools (to improve the displayed image) on PACS workstations was briefly addressed in the discussion of diagnostic accuracy. Franken (1992) related the increase in reading time on the PACS to the use of image processing software tools. Kato (1995) noted a similar time trend, but the increase in reading time in his study was not statistically significant. In their assessments of diagnostic accuracy, both Korsoff (1995) and DeCorato (1995) noted that the windowing function was not used by their readers.

Because workstation tools are designed to improve image quality, their inconsistent use may have reduced the diagnostic accuracy of workstation image interpretation reported in the literature. The impact of workstation design, as well as the adequacy of user training and support for the use of workstations, are significant unresolved issues. Support for well-designed research, as well as strategic planning for changes in organizational policy, procedures, and standards will be needed as PACS are implemented within VA.

Imaging information use by ordering clinicians: It has been assumed that clinicians use imaging information to help formulate treatment plans, and that more rapid availability of imaging information would result in more rapid clinical interventions. The findings in the 1996 Kundel study did not fully support this assumption. The study's overall finding was that, despite PACS having made imaging information available to clinicians an average of 50 minutes sooner than what had been reported for film imaging pre-PACS, this did not lead to an overall decrease in the time interval from the imaging exam to when a clinician first encountered the imaging information, except for a subset of patients. The authors suggested that high workload, established routines of work for the ICU physicians, or lack of availability of other test reports may have contributed to buffering the potential effects of PACS on improving clinical efficiency.

Although this is only one study, it is consistent with experiences in private industry. To fully achieve the potential benefits thought to be inherent in PACS, the implementation of the technology should be coordinated with the strategic re-engineering of the organization's work processes (Clyburn, 1996).

While Kundel reported no overall change in the lapsed time interval between the image exam and the access of imaging information by clinicians, he did report a major change in the way imaging information was obtained in a PACS environment versus a film-based environment. There was a significant decrease in input from radiologists when the workstations were made

available, mostly due to a significant decrease in the number of direct contacts with radiologists for imaging information. The impact of this decreased reliance on radiologists was not assessed in this study. Limited research cited in both the diagnostic accuracy evidence tables and in the literature (Levin, 1994) suggest that the accuracy of image interpretation varies widely across specialties and according to levels of training and experience.

It can be anticipated that PACS will have broad implications for the way providers practice. It can also be anticipated that organizational policies, procedures, and guidelines will need to be developed to ensure that safe standards of care are maintained as work processes change.

The paucity of empirical evidence of changes in efficiency in PACS environments, and the methodological limitations of the studies, may reflect the difficulty and costliness of designing and conducting studies of clinical and production processes in a clinical setting. However, continued efforts must be supported to assess the impact of PACS, and other medical information technologies, on processes within the health care system. Gaining a clear and comprehensive understanding of work processes is central to reorganization efforts designed to optimize system functioning.

C. ISSUE 3: What are the effects of PACS on clinical care and patient outcomes?

1. Published findings

From the articles retrieved, one study was identified which presented empirical evidence assessing the process of clinical care in a PACS environment as compared with that in a conventional film-based environment. No studies were identified which assessed the impact of PACS on patient outcomes.

Kundel (1996), as noted earlier, compared several discrete process activities in a PACS environment to those in a film-based environment, using methodologically rigorous techniques for measuring the time taken to perform the activities. The last discrete element of the process of care measured was elapsed time from the completion of an imaging exam to the performance of a clinical action based on the imaging results. Time to initiate clinical action was used as a surrogate outcome for improved quality of care.

Image-based clinical action is more likely to be necessary in a subset of patients who are critically ill, or who were recently admitted to the medical intensive care unit. For the subset of patients in which an image-based clinical action was reported and timed, it was reported that the clinical action was taken significantly faster when workstation imaging was used than when films were reviewed (2.5 hours for PACS versus 4.4 hours for film). The time span between imaging and action was then subdivided into: (a) the time interval from examination completion until the imaging information was first accessed, and (b) the time interval from access of the information until the primary action. It was determined that the shorter action time was due to a decrease in the time from exam completion until the initial access of imaging information.

2. Discussion

Imaging-based clinical actions were taken faster in a PACS environment than in a film-based environment for critically ill or recently admitted medical ICU patients. This decreased time to clinical action was used as a surrogate outcome for improved quality of care. The authors

acknowledged that a decrease in the time to take clinical action does not necessarily result in an improvement in patient outcome, and they did not assess patient outcomes directly.

The MDRC found no empirical studies showing that PACS had improved patient outcomes. Demonstrating changes in patient outcomes attributable to a diagnostic imaging technology is known to be difficult. For many diseases, the course of treatment evolves slowly, and overall patient prognosis may be determined more by patient characteristics (age, coexisting illness), rather than by early or efficient diagnosis. Therefore, demonstration of improved patient outcomes early in the development of an imaging technology is highly unlikely. However, compelling proof of the value of an innovative technology would be provided if such a demonstration could be achieved (Kent and Larson, 1992).

According to Donabedian's model of the relationships between quality, cost, and health, it is actually possible to demonstrate added value for a health care technology that produces *no improvement in patient outcomes*. This can occur if a technology has no impact on the quality of care, but does have a significant effect on increasing production efficiency. This would allow for the delivery of the current level of quality at lower cost, or for the production of more care with no change in either quality or cost.

PACS and related technologies are designed to improve both clinical and production efficiencies. Assessments of their impact on clinical and production processes, and ultimately of the cost-effectiveness, are likely to require the combined expertise of clinical and management researchers. Based on an assessment of the literature, methodologies for such research will need to be refined, and high quality research conducted.

D. ISSUE 4: What cost savings accrue from changes in processes and outcomes of care that can be attributed to the installation of a PACS?

1. Published findings

a.) Published reviews of the literature

Two published reviews of early economic evaluations of PACS were retrieved. The reviews, and several of the studies they included, could be classified as a Health Economic Research Group Stage I assessments (Table A1.5). These were early evaluations to explore the likely economic characteristics of PACS, with the overall objective of assessing the potential for PACS to be a cost-effective use of resources, should its clinical promise be realized. Since most of the studies included in the reviews were published before 1990, much of the data are now outdated. However, lessons learned about the process and utility of economic evaluations of PACS remain valuable, and could be used to help guide the next stage of economic evaluations, now that PACS technology is maturing. A brief overview of the reviews is therefore presented below.

Van Gennip et al. (1990) reviewed five economic evaluations of PACS done in the late 1980s, four of which were case reports from individual hospitals and one of which averaged data from 14 hospitals. Information reported in these studies was used to compare the costs and savings for hospital-wide PACS in the first year of operations. All costs were from the perspective of the radiology department, and costs from film-based systems were used for comparison. Annual equipment costs were calculated by adding maintenance costs to the linear amortization costs.

They reported that the annual costs of hospital-wide PACS varied between two and four million dollars. Three studies in the review reported that PACS was significantly less costly than a film-based system, with PACS costs equal to 65% - 88% of film costs in two of these

studies. Two other studies reported that PACS was 1.8 - 2.7 times more expensive than film. Differences in the findings could not be explained by variations in the sizes of the hospitals, or in the number of examinations performed. The major sources of cost variation were costs/unit, and types of personnel costs included in the analyses.

Becker and Arenson (1994) reviewed 12 economic evaluations of PACS performed at 8 sites from 1987 to 1990. Four of the studies were also included in van Gennip (1990). All costs were from the perspective of the radiology department, and costs from film-based systems were used for comparison. The authors found wide variation among studies regarding the costs and benefits of PACS, with no consistency in the conclusion as to if, or when, PACS would be less costly or more cost-effective than film. They noted that most of the PACS equipment costs were for commercial prototypes or customized systems. This precluded the ability to generalize cost findings to other settings.

b.) Published economic evaluations meeting HERG Stage II criteria

Based on the present level of maturity of the technology, the availability of many potentially evaluable PACS installations, and the existence of developed models for economic evaluation of PACS, the MDRC anticipated that well-designed studies meeting the Health Evaluation Research Group's (HERG) criteria for a Stage II economic evaluations would be identified in the recent literature.

However, no identified studies published between 1990 and 1997 met the HERG criteria for Stage II evaluation. Like the earlier studies summarized in the review articles, recent empirical studies of PACS costs used *ad hoc* methodologies and simulation. These approaches continued to be used, despite the increased availability of evaluable PACS installations, increased maturity of the technology, and increased emphasis on the need to perform high-quality economic evaluations in health care.

2. Discussion

a.) Published reviews of the literature

The two published reviews of the literature concluded that the early economic evaluations of PACS yielded conflicting results, and indicated that high quality economic analyses of PACS were still needed.

The van Gennip group developed a software package to support modeling to predict net costs of full-scale PACS implementation based on explicit criteria. Their work emphasized the need to move beyond *ad hoc* methodologies and to collect well-defined data sets using established methodologies.

Becker and Arenson suggested a need to determine whether economic evaluations, which have largely focused on direct costs, should be expanded to include indirect costs and benefits if, and when, they were empirically demonstrated. Such studies would need to broaden their perspective beyond that of the radiology department to capture possible system-wide changes related to PACS. In looking beyond the direct costs to the radiology department, Becker et al. broached a critical issue in the assessment of information technologies. The costs and the benefits of technologies that are designed to be embedded into an organization's infrastructure can be expected to extend beyond one department. True costs of ownership, and true benefits, of PACS may accrue to the organization, not just to the radiology department.

b.) Published economic evaluations meeting HERG Stage II criteria

The literature searches conducted by the MDRC demonstrated that empirical cost studies of PACS remain scarce. A critical appraisal of the articles retrieved for possible inclusion in this report concluded that recent published empirical economic evaluations have not used the stronger designs suggested in the methodologic literature, and that none of the articles met the inclusion criteria described in the Methods (Appendix 1).

The literature did indicate that higher quality studies are underway, but that their findings are not yet available. These include (but may not be limited to) ongoing efforts at both the University of Pennsylvania[1] and at Brunel University, Uxbridge[2]. Both teams are conducting a number of studies designed to explore different aspects of PACS, using several strong design and data collection methods that they have described in preliminary publications (Langlots, 1995; Keen, 1994). Both are attempting to capture changes in process and outcome from a broad perspective. The University of Pennsylvania team will also continue its ongoing assessments of the diagnostic value of PACS imaging as compared with film.

The paucity of economic evaluations of PACS, and the generally weak quality of the studies, reflect a broader problem. Quantifying costs in the health care system, and measuring changes in costs attributable to the implementation of specific technologies, have proven to be methodologically difficult (Mendelson and Salinsky, 1997). Cost accounting methods developed within traditional hospital settings have not been easily adapted to meet the needs of planners or researchers. Alternative cost accounting methods have been developed in the manufacturing sector. Two examples are Activity-Based Costing (Young and Pearlman, 1993) and Total Cost of Ownership models (Ellram, 1995, 1997). While a full discussion of these (competing) approaches is beyond the scope of this report, they are mentioned to indicate that there are tools available to health care planners and researchers that may be better suited to the task of assessing PACS costs than the methodologies with which they are now struggling.

VI. CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

An integrated information system is considered a key component of an efficiently and effectively integrated health care system. Many factors support the selection of radiology as a “service line” in which to pilot the use of information technologies to improve both clinical and production efficiency:

- radiology is costly, it is technology-intensive, and already partially computer-dependent;
- radiology data (both dictated reports and imaging) are routinely shared among providers using a mixture of hard-copy and electronic transmission;
- radiologists generally have a high level of comfort with technology, and they function largely in a consultative role with limited patient interaction.

There may be many inefficiencies in the use of human and capital resources related to medical imaging. PACS and other telemedicine technologies have the potential to fundamentally change clinical and production processes, and may support the more efficient use of resources.

---

[1] Department of Radiology and the Leonard Davis Institute of Health Economics.

[2] Health Economics Research Group, Brunel University, Uxbridge, Middlesex, U.K.

While there has been ample discussion of the *potential* of such technologies, there is a paucity of evidence demonstrating that this potential has been realized. The evidence published to date is insufficient to provide definitive answers to the critical questions posed about the use of PACS in a clinical setting. Although data do suggest that the technology is improving, it has not yet been clearly demonstrated that PACS workstation imaging is equivalent to conventional film for the accurate primary diagnosis of all of the types of illnesses that present in the veteran population. Limited available data do suggest that some work processes are performed more rapidly in a PACS environment. It remains to be demonstrated that overall clinical and production processes are more efficient, or that those efficiencies translate into improved quality, increased access, or reduced cost of care. High quality studies of effectiveness, outcomes, and cost benefit are still needed.

The data do reflect the changing nature of the technology and the complexity of implementing and assessing infrastructure change, even in a favorable environment. And they do suggest areas for further planning and research.

The Veterans Health Administration (VHA), one of the country's largest health care systems, has had a longstanding commitment to the dissemination of information technology into medicine. As a result, VHA is ahead of much of the health care delivery sector in its development of an information infrastructure. Because of its pioneering efforts to develop a health care information infrastructure, VHA now has significant resources invested in computer systems and support, and has many customized systems at individual sites to meet the needs of local users.

The integration of a medical informatics technology (PACS) into such an environment produces both opportunities and challenges that may not exist in other health care delivery systems. Organization-wide strategic planning is needed to guide VA's investments in the assessment and implementation of PACS and telemedicine.

Suggested areas for further planning and research:

- **clarify and disseminate** strategic goals and objectives for VA telemedicine activities.
- **strengthen the oversight and coordination** of PACS and other telemedicine activities within VA and among federal agencies. The design and implementation of coordinated multi-site assessments of PACS should be included in these efforts.
- further integrate management and economic research methods into PACS assessments to better capture the costs and effects. The literature suggests that we are beginning to understand how to study information technologies, and some useful techniques have been developed in the fields of economics and organizational research. Implicit in this is the need to **strengthen the working relationships among health care economists, management researchers, and clinical researchers within VA.**
- **practice evidence-based decision-making and purchasing**. Require (and fund) more rigorous evaluation of effectiveness, and cost-effectiveness, before broadly disseminating PACS. While there is much enthusiasm for PACS expressed by proponents, there is limited evidence on which to base a recommendation for wide dissemination at this time.
- **assess information technologies from an institutional perspective**. Many cost evaluations reflect the traditional focus of information systems and capital expenses on individual departments (usually radiology). In an integrated health care delivery system, the costs and potential benefits associated with PACS, and with the information infrastructure upgrades needed for their implementation, are likely to be system-wide.
- continue to **support test beds for the refinement and evaluation of commercial PACS and in-house image management systems**. While the technology appears to be well-aligned with the mission and strategic objectives of our health care system, credible empirical evidence of efficiency, effectiveness, or economic benefit are still needed.

ADDENDUM

Since this report was written, a paper was submitted for publication by the Leonard Davis Institute of Health Economics and Department of Radiology, University of Pennsylvania. The report, entitled *The Incremental Cost of a Department-Wide PACS/CR Implementation*, is scheduled for publication early next year.

References

REFERENCES

The following articles met inclusion criteria for this report, or were cited to support key points in the report .

Ackerman SJ, Gitlin JN, Gayler RW, Flagle CD, Bryan RN. Receiver operating characteristic analysis of fracture and pneumonia detection: comparison of laser-digitized workstation images and conventional analog radiographs. *Radiology* 1993; 186(1): 263-8.

Advisory Committee on Telecommunications and Health Care. (1996). *FCC telecommunications and health care advisory committee: findings and recommendations* [website]. Available: http://www.fcc.gov/Reports/hc071196.txt [April 1997].

American College of Radiology. (1996). *ACR standard for teleradiology (Res 26-1996)* [website]. Available: http://www.acr.org [February, 1997].

Bakker AR, Stut WJ, De Valk JP, Reijns GR. PACS costs: modeling and simulation. *Medical Informatics* 1988; 13(4): 307-13.

Banta HD, Luce BR. Picture archiving and communication systems (PACS): an example of the computer revolution. In: Banta H, Luce B, eds. *Health Care Technology and its Assessment: An International Perpsective.* New York: Oxford University Press; 1993.

Becker SH, Arenson RL. Costs and benefits of picture archiving and communication systems. *Journal of the American Medical Informatics Association* 1994; 1(5): 361-71.

Bradham DD, Morgan S, Dailey ME. The information superhighway and telemedicine: applications, status, and issues. *Wake Forest Law Review* 1995; 30(1): 145-66.

Center for The Evaluation of Healthcare Information Systems Technology. *Digital imaging technology assessment: final report.* 1994, Medical Information Resources Management Office. Department of Veterans Affairs: Martinsburg.

Clyburn CA, Department of Defense. Telemedicine Test Bed. (1996). *An integrated process re-engineering model for improving health care delivery using telemedicine* [website]. Available: http://206.156.10.15/pages/library/papers/reengineer/procreng.html [December, 1996].

Cox GG, Cook LT, McMillan JH, Rosenthal SJ, Dwyer SJ. Chest radiography: comparison of high-resolution digital displays with conventional and digital film. *Radiology* 1990; 176(3): 771-6.

Dayhoff R. (1997). *VISTA medical imaging update* [website]. Available: http://www.va.gov/vama.htm [July, 1997].

DeCorato DR, Kagestu NJ, Ablow RC. Off-hours interpretation of radiologic images of patients admited to the emergency department: efficacy of teleradiology. *AJR: American Journal of Roentgenology* 1995; 165(5): 1293-96.

Deloitte & Touche Consulting Group. (1997). *Technical workstation total cost of ownership study: a cost comparison of Microsoft Windows NT Workstation and UNIX* [website]. Available: http://www.alphastation.digital.com/announce/tco/tco an.html [March, 1997].

Department of Health & Human Services. Health Care Financing Administration. (1997). *Telemedicine services* [website]. Available: http://www.hcfa.gov/medicare/op1041.htm [March 1, 1997].

DeSimone DN, Kundel HL, Arenson RL, Seshadri SB, Brikman IS, Khalsa SS, Davey MJ, Brisbon NE. Effect of a digital imaging network on physician behavior in an intensive care unit. *Radiology* 1988; 169(1): 41-4.

Donabedian A, Wheeler J, Wyszewianski L. Quality, cost, and health: an integrative model. *Medical Care* 1982; 20(10): 975-91.

Elam EA, Rehm K, Hillman BJ, Maloney K, Fajardo LL, McNeill K. Efficacy of digital radiography for the detection of pneumothorax: comparison with conventional chest radiography. *AJR: American Journal of Roentgenology* 1992; 158(3): 509-14.

Ellram LM. A taxonomy of total cost of ownership models. *Journal of Business Logistics* 1994; 15(1): 171-91.

Ellram LM. Total cost of ownership: an analysis approach for purchasing. *International Journal of Physical Distribution & Logistics Management* 1995; 25(8): 4-23.

Enning CJ, van Gennip EM, van den Broeck R, Rechid R, Wein B, Bakker AR. PACER: a software tool for PACS decision makers. *Medical Informatics* 1994; 19(2): 179-87.

Eysenck HJ. Meta-analysis and its problems. *BMJ* 1994; 309(6957): 789-792.

Flynn K, Adams E. *Assessing diagnostic technologies.* MTA94-001-01. 1996, Department of Veterans Affairs. Health Services Research & Development Service. Management Decision & Research Center: Boston.

Franken EA, Jr., Berbaum KS, Marley SM, Smith WL, Sato Y, Kao SC, Milam SG. Evaluation of a digital workstation for interpreting neonatal examinations. A receiver operating characteristic study. *Investigative Radiology* 1992; 27(9): 732-7.

Gay SB, Sobel AH, Young LQ, Dwyer SJ, III. Processes involved in reading imaging studies: workflow analysis and implications for workstation development. *Journal of Digitial Imaging* 1997; 10(1): 40-5.

Glass HI. The impact of PACS on the information flow. *International Journal of Bio-Medical Computing* 1992; 30(3-4): 229-34.

Goldberg MA, Belanger G, Tombaugh J, Dillon R, Robertson J, Corstine M, Mastronardi J, Beeton C. Clinical experience with a multimedia workstation at the Ottawa Civic Hospital, Canada. In: Huang H, ed. *Picture Archiving and Communication Systems (PACS) in Medicine.* Berlin: Springer-Verlag; 1991.

Goldberg MA, Rosenthal DI, Chew FS, Blickman JG, Miller SW, Mueller PR. New high-resolution teleradiology system: prospective study of diagnostic accuracy in 685 transmitted clinical cases. *Radiology* 1993; 186(2): 429-34.

Guyatt G, Sackett D, Sinclair J, Hayward R, Cook D, Cook R. User's guide to the medical literature: IX. A method for grading health care recommendations. *JAMA* 1995; 274: 1800-4.

Haynes RB, Sackett D. Purpose and procedure (abbreviated). *Evidence-Based Medicine* 1995; 1(1): 2.

Heymann TD, Culling W. Ward based x-ray facilities can improve services. *Journal of the Royal College of Physicians of London* 1996; 30(2): 145-9.

Honeyman J, Staab E, Frost M, Brandt T, Bidgood W. Tools for analysis of requirements for PACS. In: Huang H, ed. *Picture Archiving and Communication Systems (PACS) in Medicine.* Berlin: Springer-Verlag; 1991.

Horii SC, Levine BA, Goger G, Mun SK, Fielding R, Garra BS, Lo SC, Krasner B, Benson H. A comparison of case retrieval times: film versus picture archiving and communications systems. *Journal of Digital Imaging* 1992; 5(3): 138-43.

Ibrahim MA. *Epidemiology and health policy.* Gaithersburg: Aspen Systems Corporation; 1985.

Kato H, Kubota G, Kojima K, Hayashi N, Nishihara E, Kura H, Aizawa M. Preliminary time-flow study: comparison of interpretation times between PACS workstations and films. *Computerized Medical Imaging & Graphics* 1995; 19(3): 261-5.

Keen J, Bryan S, Muris N, Weatherburn G, Buxton M. A Model for the Evaluation of PACS. In: Boehme J, Rowberg A, Wolfman N, eds. *SCAR 94 Computer Applications to Assist Radiology.* Boston: Symposia Foundation; 1994.

Kent DL, Haynor DR, Longstreth WT, Jr., Larson EB. The clinical efficacy of magnetic resonance imaging in neuroimaging. *Annals of Internal Medicine* 1994; 120(10): 856-71.

Kent DL, Larson EB. Magnetic resonance imaging of the brain and spine. Is clinical efficacy established after the first decade? *Annals of Internal Medicine* 1988; 108(3): 402-24.

Kent DL, Larson EB. Disease, level of impact, and quality of research methods. Three dimensions of clinical efficacy assessment applied to magnetic resonance imaging. *Investigative Radiology* 1992; 27(3): 245-54.

Kizer KW. *Prescription for change: the guiding principles and strategic objectives underlying the transformation of the Veterans healthcare system.* Washington, DC: U.S. Government Printing Office; 1996.

Kondoh H, Ikezoe J, Inamura K, Kuroda C, Kozuka T. A comparison of conventional screen-film radiography and hard copy of computed radiography in full and two-thirds sizes in detection of interstitial lung disease. *Journal of Digital Imaging* 1994; 7(4): 193-5.

Korsoff L, Kallio T, Kormano M, Heinila J. Experiences with a teleradiology system in pulmonary diseases. *Acta Radiologica* 1995; 36(1): 37-40.

Kundel HL. Visual perception and image display terminals. *Radiologic Clinics of North America* 1986; 24(1): 69-78.

Kundel HL, Seshadri SB, Langlotz CP, Lanken PN, Horii SC, Nodine CF, Polansky M, Feingold E, Brikman I, Bozzo M, Redfern R. A prospective study of PACS - information flow and clinical action in a medical intensive care unit. *Radiology* 1996; 199(1): 143-9.

Langlotz CP, Even-Shoshan O, Seshadri SS, Brikman I, Kishore S, Kundel HL, Schwartz JS. A methodology for the economic assessment of picture archiving and communication systems. *Journal of Digital Imaging* 1995; 8(2): 95-102.

Levin DC. Merrill C. Sosman Lecture. The practice of radiology by nonradiologists: cost, quality, and utilization issues. *AJR: American Journal of Roentgenology* 1994; 162(3): 513-8.

Lou SL, Huang HK. Assessment of a neuroradiology picture archiving and communication system in clinical practice. *AJR: American Journal of Roentgenology* 1992; 159(6): 1321-7.

Luce BR, Simpson K. Methods of cost-effectiveness analysis: areas of consensus and debate. *Clinical Therapeutics* 1995; 17(1): 109-25.

Mendelson DN, Salinsky EM. Health information systems and the role of state government. *Health Affairs* 1997; 16(3): 106-19.

Mulrow CD. Rationale for systematic reviews. *BMJ* 1994; 309(6954): 597-99.

Mulrow CD. The medical review article: state of the science. *Annals of Internal Medicine* 1987; 106(3): 485-8.

Paakkala T, Aalto J, Kahara V, Seppanen S. Diagnostic performance of a teleradiology system in primary health care. *Computer Methods & Programs in Biomedicine* 1991; 36(2-3): 157-60.

Persson J, Ekberg K, Linden M. Work organisation, work environment and the use of medical equipment: a survey study of the impact on quality and safety. *Medical & Biological Engineering & Computing* 1993; 31(1): HTA20-4.

Petitti DB. *Meta-anaylsis, decision making, and cost-efectiveness analysis: methods for quantitative sysnthesis in medicine.* New York: Oxford University Press; 1994.

Pocock SJ. *Clinical trials: a practical approach.* Chichester: John Wiley & Sons; 1991.

Razavi M, Sayre JW, Taira RK, Simons M, Huang HK, Chuang KS, Rahbar G, Kangarloo H. Receiver-operating-characteristic study of chest radiographs in children: digital hard-copy film vs 2K x 2K soft-copy images. *AJR: American Journal of Roentgenology* 1992; 158(2): 443-8.

Rigby K, Silagy C, Crockett A. Health economic reviews. Are they compiled systematically? *International Journal of Technology Assessment in Health Care* 1996; 12(3): 450-9.

Scott WW, Jr., Bluemke DA, Mysko WK, Weller GE, Kelen GD, Reichle RL, Weller JC, Gitlin JN. Interpretation of emergency department radiographs by radiologists and emergency medicine physicians: teleradiology workstation versus radiograph readings. *Radiology* 1995; 195(1): 223-9.

Scott WWJ, Rosenbaum JE, Ackerman SJ, Reichle RL, Magid D, Weller JC, Gitlin JN. Subtle orthopedic fractures - teleradiology workstation versus film interpretation. *Radiology* 1993; 187(3): 811-5.

Sculpher M, Drummond M, Buxton M. *Economic evaluation in health care research and development: undertake it early and often.* Number 12. 1995, HERG Discussion Paper: Uxbridge.

Slasky BS, Gur D, Good W, Costa-Greco M, Harris K, Cooperstein L, Rockette H. Receiver operating characteristic analysis of chest image interpretation with conventional, laser-printed, and high-resolution workstation images. *Radiology* 1990; 174(3 part 1): 775-80.

Starr P. Smart technology, stunted policy: developing health information networks. *Health Affairs* 1997; 16(3): 91-105.

Straub WH, Gur D, Good WF, Campbell WL, Davis PL, Hecht ST, Skolnick ML, Thaete FL, Rosenthal MS, Sashin D. Primary CT diagnosis of abdominal masses in a PACS environment. *Radiology* 1991; 178(3): 739-43.

Thaete FL, Fuhrman CR, Oliver JH, Britton CA, Campbell WL, Feist JH, Straub WH, Davis PL, Plunkett MB. Digital radiography and conventional imaging of the chest: a comparison of observer performance. *AJR: American Journal of Roentgenology* 1994; 162(3): 575-81.

Tohme WG, Olsson S. Developments and directions of technology assessment in telemedicine. *Telemedicine: new articles published in 1995.* Stockholm: Swedish Institute for Health Services Development. SPRI; 1996.

Tucker DM, Barnes GT, Koehler RE. Picture archiving communication systems in the intensive care unit. *Radiology* 1995; 196(2): 297-304.

U. S. Department of Commerce. Joint Working Group on Telemedicine. (1997). *Telemedicine report to Congress* [website]. Available: http://www.ntia.doc.gov/reports/telemed/index.htm [January 31, 1997].

van Gennip EMSJ, Bakker AR. Assessment of effects and costs of information systems. *International Journal of Bio-Medical Computing* 1995; 39(1): 67-72.

van Gennip EMSJ, Ottes FP, van Poppel BM, Andriessen JHTH. Why do cost-benefit studies of PACS disagree? *SPIE* 1990; 1234: 894-904.

Warburton RN. Evaluation of PACS-induced organizational change. *International Journal of Bio-Medical Computing* 1992; 30(3-4): 243-8.

Warburton RN, Fisher PD, Nosil J, Brauer GW, Lawrence WJ, Ritchie GW. Digital diagnostic imaging with a comprehensive PACS: hypothetical economic evaluation at a large community hospital. *Journal of Digital Imaging* 1990; 3(2): 101-7.

Washowich TL, Williams SC, Richardson LA, Simmons GE, Dao NV, Allen TW, Hammet GC, Morris MJ. Detection of interstitial lung abnormalities on picture archive and communication system video monitors. *Journal of Digital Imaging* 1997; 10(1): 34-9.

Yoshino MT, Carmody R, Fajardo LL, Seeger J, Jones K. Diagnostic performance of teleradiology in cervical spine fracture detection. *Investigative Radiology* 1992; 27(1): 55-9.

Young DW, Pearlman LK. Managing the stages of hospital cost accounting. *Healthcare Financial Management* 1993; 47(4): 58, 60, 63-4 passim.

Selected Bibliography

SELECTED BIBLIOGRAPHY
The articles below address significant content or methodologic issues, but do not necessarily meet the inclusion criteria for this report.

Berger SB, Cepelewicz BB. Medical-legal issues in teleradiology. *AJR: American Journal of Roentgenology* 1996; 166(3): 505-10.

Boehme JM, Choplin RH. PACS mini refresher course. Systems integration: requirements for a fully functioning electronic radiology department. *RadioGraphics* 1992; 12(4): 789-94.

Canadian Coordinating Office for Health Technology Assessment. *Guidelines for economic evaluation of pharmaceuticals: Canada.* 1994, Canadian Coordinating Office for Health Technology Assessment: Ottawa.

Chapko MK, Ehreth JL, Hedrick S. Methods of determining the cost of health care in department of veterans affairs medical centers and other non-priced settings. *Evaluation & the Health Professions* 1991; 14(3): 282-303.

Cook DJ, Guyatt GH, Laupacis A, Sackett D. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest* 1992; 102(4, Supplement): 305S-11S.

Cooper R, Kaplan RS. Measure costs right: make the right decisions. *Harvard Business Review* 1988; 88(1): 96-103.

Crowe BL. Overview of some methodological problems in assessment of PACS. *International Journal of Bio-Medical Computing* 1992; 30(3-4): 181-6.

Department of Clinical Epidemiology and Biostatistics. McMaster University Health Sciences Centre. How to read clinical journals. VII. To understand an economic evaluation (part B). *Canadian Medical Association Journal* 1984; 130(12): 1542-49.

Department of Clinical Epidemiology and Biostatistics. McMaster University Health Sciences Centre. How to read clinical journals. VII. To understand an economic evaluation (part A). *Canadian Medical Association Journal* 1984; 130(11): 1428-34.

Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996; 313(7052): 275-83.

Drummond MF, Soddart GL, Torrance G. *Methods for the economic evaluation of health care programs.* New York: Oxford Medical Publications; 1987.

Dwyer SJ, Stewart BK, Sayre JW, Aberle DR, Boechat MI, Honeyman JC, et al. PACS mini refresher course. Performance characteristics and image fidelity of gray-scale monitors. *RadioGraphics* 1992; 12(4): 765-772.

Eddy DM. *A manual for assessing health practices & designing practice policies: the explicit approach.* Philadelphia: American College of Physicians; 1992.

Field MJ. *Telemedicine: a guide to assessing telecommunications in health care.* Washington DC: National Academy Press; 1996.

Forman HP, McClennan BL. Meeting of managerial science with medicine. The pace quickens. *JAMA* 1996; 276(19): 1599-1600.

Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Medical Decision Making* 1991; 11(2): 88-94.

Goodman C. *Literature searching and evidence interpretation for assessing health care practices.* Stockholm: Swedish Council on Technology Assessment in Health Care; 1993.

Grigsby J, Schlenker RE, Kaehny MM, Shaughnessy PW, Sandberg EJ. Analytic framework for evaluation of telemedicine. *Telemedicine Journal* 1995; 1(1): 31-39.

Honeyman JC, Messinger JM, Frost MM, Staab EV. PACS mini refresher course. Evaluation of requirements and planning for picture archiving and communication systems. *RadioGraphics* 1992; 12(1): 141-50.

Horii SC. PACS mini refresher course. Electronic imaging workstations: ergonomic issues and the user interface. *RadioGraphics* 1992; 12(4): 773-87.

Huang HK. PACS mini refresher course. Three methods of implementing a picture archiving and communication system. *RadioGraphics* 1992; 12(1): 131-39.

Irwig L, Tosteson AN, Gatsonis C, Lau J, Colditz G, Chalmers TC, et al. Guidelines for meta-analyses evaluating diagnostic tests. *Annals of Internal Medicine* 1994; 120(8): 667-76.

Keen J, Bryan S, Muris N, Weatherburn G, Buxton M. Evaluation of diffuse technologies: the case of digital imaging networks. *Health Policy* 1995; 34(3): 153-66.

Langlotz CP, Seshadri S. Technology assessment methods for radiology systems. *Radiologic Clinics of North America* 1996; 34(3): 667-79.

Metz CE. Some practical issues of experimental design and data analysis in radiological ROC studies. *Investigative Radiology* 1989; 24(3): 234-45.

Mulrow CD, Linn WD, Gaul MK, Pugh JA. Assessing quality of a diagnostic test evaluation. *Journal of General Internal Medicine* 1989; 4(4): 288-95.

Mun SK, Benson H, Welsh C, Elliot L, Davros W. Baseline study of radiology services for the purpose of PACS evalution. *SPIE Medical Imaging II* 1988; 914: 978-87.

Riegelman RK, Hirsch RP. *Studying a study and testing a test: how to read the health science literature.* Boston: Little Brown; 1996.

Sheldon TA. Problems of using modelling in the economic evaluation of health care. *Health Economics* 1996; 5(1): 1-11.

Siegel E, Allman R, Gitlin J, Pickar E. Parameters required to evaluate a filmless imaging system. In: Boehme J, Rowberg A, Wolfman N, eds. *SCAR 94 Computer Applications to Assist Radiology.* Boston, MA: Symposia Foundation; 1994.

Siegel EL, Diaconis JN, Pomerantz S, Allman R, Briscoe B. Making filmless radiology work. *Journal of Digital Imaging* 1995; 8(4): 151-55.

Task Force on Principles for Economic Analysis of Health Care Technology . Economic analysis of health care technology. A report on principles. *Annals of Internal Medicine* 1995; 123(1): 61-70.

US General Accounting Office Report to Congressional Requestors. *Telemedicine: Federal strategy is needed to guide investments.* GAO/NSAID/HEHS-97-67. 1997, US General Accounting Office: Washington, DC.

van Gennip EMSJ, Bakker AR. Challenges and opportunities for technology assessment in medical informatics. Case study: PACS. *Medical Informatics* 1993; 18(3): 209-18.

van Gennip EMSJ, Bakker AR, Greberman M. Do the benefits outweigh the costs of PACS? The results of an international workshop on technology assessment of PACS. *Computer Methods & Programs in Biomedicine* 1992; 37(4): 265-71.

van Gennip EMSJ, Talmon JL. *Assessment and evaluation of information technologies in medicine.* Amsterdam, Washington DC: IOS Press; 1995.

Ineligible Studies

INELIGIBLE STUDIES
The following articles were reviewed, but failed to meet the inclusion criteria.

Aberle DR, Gleeson F, Sayre JW, Brown K, Batra P, Young DA, et al. The effect of irreversible image compression on diagnostic accuracy in thoracic imaging. *Investigative Radiology* 1993; 28(5): 398-403.

Arenson RL, Burnside ES, Avrin DE, Gould RG, Huang HK, Marco RP. Cost benefits of picture archiving and communications systems. *Academic Radiology* 1996; 3(Suppl 1): S 72-S 74.

Arenson RL, Seshadri SB, Kundel HL, DeSimone D, Van der Voorde F, Gefter WB, et al. Clinical evaluation of a medical image management system for chest images. *AJR: American Journal of Roentgenology* 1988; 150(1): 55-9.

Bakker AR, Stut WJ, De Valk JP, Reijns GR. PACS costs: modeling and simulation. *Medical Informatics* 1988; 13(4): 307-13.

Bauman RA. Worldwide experience with large PACS systems. In: Boehma J, Rowberg A, Wolfman N, eds. *Computer Applications to Assist Radiology*. Boston: Symposia Foundation; 1994.

Bauman RA, Gell G, Dwyer SJ. Large picture archiving and communication systems of the world. 1. *Journal of Digital Imaging* 1996; 9(3): 99-103.

Beard DV, Hemminger BM, Perry JR, Mauro MA, Muller KE, Warshauer DM, et al. Interpretation of CT studies: single-screen workstation versus film alternator. *Radiology* 1993; 187(2): 565-9.

Beard DV, Molina PL, Muller KE, Denelsbeck KM, Hemminger BM, Perry JR, et al. Interpretation time of serial chest CT examinations with stacked-metaphor workstation versus film alternator. *Radiology* 1995; 197(3): 753-8.

Beard DV, Parrish D, Stevenson D. A cost analysis of film image management and four PACS systems. *SPIE Medical Imaging IV: System Design and Evaluation* 1990; 1234: 10-17.

Brill PW, Winchester P, Cahill P, Lesser M, Durfee SM, Giess CS, et al. Computed radiography in neonatal and pediatric intensive care units: a comparison of 2.5 K x 2 K soft-copy images vs digital hard-copy film. *Pediatric Radiology* 1996; 26(5): 333-6.

Cannavo MJ. PACS prices, performance showing improvement. *Health Management Technology* 1996; 17(2): 22, 24, 44.

Cawthon MA, Mogel GT, Williamson MP. Soft copy interpretation of emergency department examinations. *Emergency Radiology* 1996; 3(1): 30-3.

Chakraborty DP, Breatnach ES, Yester MV, Soto B, T. BG, Fraser RG. Digital and conventional chest imaging: a modified ROC study of observer performance using simulated nodules. *Radiology* 1986; 158(1): 35-9.

Dershaw DD, Fleischman RC, Liberman L, Deutch B, Abramson AF, Hann L. Use of digital mammography in needle localization procedures. *AJR: American Journal of Roentgenology* 1993; 161(3): 559-62.

DeSimone DN, Kundel HL, Arenson RL, Seshadri SB, Brikman IS, Khalsa SS, et al. Effect of a digital imaging network on physician behavior in an intensive care unit. *Radiology* 1988; 169(1): 41-4.

Don S, Albertina MJ, Ammann DL, Evens RG, Siegel MJ. Soft-copy computed radiography in neonatal and pediatric intensive care units: cost-savings analysis. *Radiology* 1995; 197(2): 501-5.

Donnelly LF, Johnson ND, Taylor CN. Increased efficiency of radiology information management with a radiology support system. *AJR: American Journal of Roentgenology* 1997; 168(3): 611-2.

Drew PG. Justifying PACS. *Health Management Technology* 1997; 18(2): 84, 86, 88.

Duerinckx AJ, Grant EG, Hagan GT, Wanchoo V. Medical experience with an ultrasound PACS at the VA Medical Center, West Los Angeles. In: Boehme J, Rowberg A, Wolfman N, eds. *SCAR 94 Computer Applications to Assist Radiology*. Boston: Symposia Foundation; 1994.

Duerinckx AJ, Valentino DJ, Hayrapetian A, Hagan G, Grant EG. Ultrafast networks (ATM): first clinical experiences. *European Journal of Radiology* 1996; 22(3): 186-96.

Enning CJ, van Gennip EM, van den Broeck R, Rechid R, Wein B, Bakker AR. PACER: a software tool for PACS decision makers. *Medical Informatics* 1994; 19(2): 179-87.

Fajardo LL, Hillman BJ, Pond GD, Carmody RF, Johnson JE, Ferrell WR. Detection of pneumothorax: comparison of digital and conventional chest imaging. *AJR: American Journal of Roentgenology* 1989; 152(3): 475-80.

Fajardo LL, Yoshino M, Seeley G, Hunt R, Hunter T, Friedman R, et al. Detection of breast abnormalities on teleradiology transmitted mammograms. *Investigative Radiology* 1990; 25(10): 1111-15.

Frank MS, Jost RG, Molina PL, Anderson DJ, Solomon SL, Whitman RA, et al. High-resolution computer display of portable, digital, chest radiographs of adults: suitability for primary interpretation. *AJR: American Journal of Roentgenology* 1993; 160(3): 473-7.

Franken EA, Jr., Smith WL, Berbaum KS, Kao SC, Sato Y. Comparison of a PACS workstation with conventional film for interpretation of neonatal examinations: a paired comparison study. *Pediatric Radiology* 1991; 21(5): 336-40.

Freedman M, Steller D. Digital radiography of the musculoskeletal system: the optimal image. *Journal of Digital Imaging* 1995; 8(1 Suppl 1): 37-42.

Halpern EJ, Newhouse JH, Amis ES, Jr., Lubetsky HW, Jaffe RM, Esser PD, et al. Evaluation of teleradiology for interpretation of intravenous urograms. *Journal of Digital Imaging* 1992; 5(2): 101-6.

Hilsenrath PE, Smith WL, Berbaum KS, Franken EA, Owen DA. Analysis of the cost-effectiveness of PACS. *AJR: American Journal of Roentgenology* 1991; 156(1): 177-80.

Hirota H, Shimamoto K, Yamakawa K, Ishigaki T, Takahashi Y, Sugiyama N, et al. Clinical evaluation of newly developed CRT viewing station - CT reading and observers performance. *Computerized Medical Imaging & Graphics* 1995; 19(3): 281-5.

Holman BL, Aliabadi P, Silverman SG, Weissman BN, Rudolph LE, Fener EF. Medical impact of unedited preliminary radiology reports. *Radiology* 1994; 191(2): 519-21.

Horii SC, Mun SK, Elliott LP, Levine B, Lo B, Garra BS. PACS clinical experience at Georgetown University. *NATO ASI Series* 1991; 74: 170-88.

Horii SC, Mun SK, Elliott LP, Levine B, Lo B, Garra BS, et al. PACS clinical experience at Georgetown University. *International Journal of Bio-Medical Computing* 1992; 30(3-4): 275-80.

Huda W, Honeyman JC, Frost MM, Staab EV. A cost-analysis of computed radiography and picture archiving and communication systems in portable radiography. *Journal of Digital Imaging* 1996; 9(1): 39-44.

Huda W, Honeyman JC, Palmer CK, Frost MM, Staab EV. Computed radiography and film digitizer inputs to an intensive care unit teleradiology system: an image quality comparison. *Academic Radiology* 1996; 3(2): 110-4.

Humphrey LM, Fitzpatrick K, Atallah N, Ravin CE. Time comparison of intensive care units with and without digital viewing systems. *Journal of Digital Imaging* 1993; 6(1): 37-41.

Inamura K, Umeda T, Harauchi H, Kondoh H, Hasegawa T, Kozuka T, et al. Time and flow study results before and after installation of a hospital information system and radiology information system and before clinical use of a picture archiving and communication system. *Journal of Digital Imaging* 1997; 10(1): 1-9.

Irie G, Miyasaka K. Clinical Experience - 16 Months of HU-PACS. *Computerized Medical Imaging and Graphics* 1991; 15(3): 191-95.

Ishida M, Frank PH, Doi K, Lehr JL. High quality digital radiographic images: improved detection of low-contrast objects and preliminary clinical studies. *Radiographics* 1983; 3(2): 325-38.

Joergensen J, Lundbak B, Egeblad M. Clinical experience with PACS. *SPIE Medical Imaging 1994 PACS: Design and Evaluation* 1994; 2165: 763-71.

Kangarloo H. PACS--clinical experience at UCLA. *Computerized Medical Imaging & Graphics* 1991; 15(3): 201-3.

Keen J, Bryan S, Muris N, Weatherburn G, Buxton M. A Model for the Evaluation of PACS. In: Boehme J, Rowberg A, Wolfman N, eds. *SCAR 94 Computer Applications to Assist Radiology.* Boston: Symposia Foundation; 1994.

Krupinski EA, Lund PJ. Differences in time to interpretation for evaluation of bone radiographs with monitor and film viewing. *Academic Radiology* 1997; 4(3): 177-82.

Kundel H, Seshardi SB, Arenson R. Clinical experience with PACS at the University of Pennsylvania. *Computerized Medical Imaging and Graphics* 1991; 15(3): 197-200.

Kundel HL, Seshadri SB, Langlotz C, Kishore S, Brikman I, Purcell JA, et al. Strategies for the clinical evaluation of PACS: experience with experimental design in a medical intensive care unit. In: Boehme J, Rowberg A, Wolfman N, eds. *Computer Applications to Assist Radiology.* Boston: Symposia Foundation; 1994.

Kurdziel JC, Kanz R. Administrative and organization impact of PACS and its effect on the quality of care. *International Journal of Bio-Medical Computing* 1992; 30(3-4): 259-62.

Lams PM, Cocklin ML. Spatial resolution requirements for digital chest radiographs: an ROC study of observer performance in selected cases. *Radiology* 1986; 158(1): 11-19.

Langlotz CP, Even-Shoshan O, Seshadri SS, Brikman I, Kishore S, Kundel HL, et al. A methodology for the economic assessment of picture archiving and communication systems. *Journal of Digital Imaging* 1995; 8(2): 95-102.

Lyttkens K, Andersson B, Kehler M, Lindberg L. Bedside chest radiography using digital luminescence. A comparison between digital radiographs reviewed on a personal computer and as hard-copies. *Acta Radiologica* 1992; 33(5): 427-30.

MacMahon H. Digital chest radiography at the University of Chicago: present status and future plans. *Journal of Digital Imaging* 1995; 8(1 Suppl 1): 11-4.

MacMahon H, Metz CE, Doi K, Kim T, Giger ML, Chan H. Digital chest radiography: effect on diagnostic accuracy of hard copy, conventional video and reversed gray scale video display formats. *Radiology* 1988; 168(3): 669-73.

Mascarini C, Ratib O, Trayser G, Ligier Y, Appel RD. In-house access to PACS images and related data through World Wide Web. *European Journal of Radiology* 1996; 22(3): 218-20.

Mosser H, Urban M, Partan G, Hruby W. Clinical routine operation of a filmless radiology department: 20 months experience. In: Boehme J, Rowberg A, Wolfman N, eds. *SCAR 94 Computer Applications to Assist Radiology.* Boston: Symposia Foundation; 1994.

Mun SK, Benson H, Welsh C, Elliot L, Davros W. Baseline study of radiology services for the purpose of PACS evalution. *SPIE Medical Imaging II* 1988; 914: 978-87.

Murphey MD, Bramble JM, Cook LT, Martin NL, Dwyer SJD. Nondisplaced fractures: spatial resolution requirements for detection with digital skeletal imaging. *Radiology* 1990; 174(3 Pt 1): 865-70.

Neumann PJ, Parente ST, Paramore LC. Potential savings from using information technology applications in health care in the United States. *International Journal of Technology Assessment in Health Care* 1996; 12(3): 425-35.

Nishihara E, Kura H, Kubota G, Kohda T. Control method for preloading with priority information in an integrated radiology information system/picture archiving and communication system. *Journal of Digital Imaging* 1997; 10(1): 27-33.

Perednia DA, Gaines JA, Butruille TW. Comparison of the clinical informativeness of photographs and digital imaging media with multiple-choice receiver operating characteristic analysis. *Archives of Dermatology* 1995; 131(3): 292-7.

Protopapas Z, Siegel EL, Reiner BI, Pomerantz SM, Pickar ER, Wilson M, et al. Picture archiving and communication system training for physicians - lessons learned at the Baltimore VA Medical Center. *Journal of Digital Imaging* 1996; 9(3): 131-6.

Richmond R, Piriano D, Ross J, Rueger W. Physician acceptance of primary intrepretation of CT and MR examinations on a workstation: implications for future functional design. In: Boehme J, Rowberg A, Wolfman N, eds. *SCAR 94 Computer Applications to Assist Radiology.* Boston: Symposia Foundation; 1994.

Sheline ME, Brikman I, Epstein DM, Mezrich JL, Kundel HL, Arenson RL. The diagnosis of pulmonary nodules: comparison between standard and inverse digitized images and conventional chest radiographs. *AJR American Journal of Roentgenology* 1989; 152(2): 261-3.

Siegel EL, Diaconis JN, Pomerantz S, Allman R, Briscoe B. Making filmless radiology work. *Journal of Digital Imaging* 1995; 8(4): 151-55.

Siegel EL, Protopapas Z, Reiner B, Pomerantz S, Cameron EW, Pickar E. A prospective evaluation of the impact of filmless operation on the Baltimore VA Medical Center. *RBM Revue Europeenne de Technologie Biomedicale* 1996; 18(5): 149-152.

Smith DV, Smith S, Bender GN, Carter JR. *Lessons learned and two years clinical experience in implementing the medical diagnostic imaging support (MDIS) system at Madigan Army Medical Center. DoD Telemedicine Test Bed* [website]. Available: http://206.156.10.15/pages/library/papers/mdisPapers/spie94mamc.html [June 24, 1997].

Steckel RJ, Batra P, Johnson S, Sayre J, Brown K, Haker K, et al. Comparison of hard- and soft-copy digital chest images with different matrix sizes for managing coronary care unit patients. *AJR: American Journal of Roentgenology* 1995; 164(4): 837-41.

Stockburger WT, King WE. PACS: a financial analysis for economic viability. *Applied Radiology* 1990; 19(1): 17-24.

Straub WH, Gur D. The hidden costs of delayed access to diagnostic imaging information: impact on PACS implementation. *AJR: American Journal of Roentgenology* 1990; 155(3): 613-6.

Strickland NH. Some cost-benefit considerations for PACS: a radiological perspective. *British Journal of Radiology* 1996; 69(828): 1089-98.

Strickland NH, Allison DJ. Default display arrangements of images on PACs monitors. *British Journal of Radiology* 1995; 68(807): 252-60.

Terae S, Hawkin S, Sato Y, Kikuchi Y, Abe S, Miyasaka K. Clinical experience of Hokkaido University-PACS and FCR-angiography. *Computer Methods & Programs in Biomedicine* 1994; 43(1-2): 93-100.

van der Loo RP, van Gennip EM. Evaluation of personnel savings through PACS: a modelling approach. *International Journal of Bio-Medical Computing* 1992; 30(3-4): 235-41.

van Gennip EMSJ, Bakker AR. Challenges and opportunities for technology assessment in medical informatics. Case study: PACS. *Medical Informatics* 1993; 18(3): 209-18.

van Gennip EMSJ, Bakker AR. Assessment of effects and costs of information systems. *International Journal of Bio-Medical Computing* 1995; 39(1): 67-72.

van Gennip EMSJ, Bakker AR, Greberman M. Do the benefits outweigh the costs of PACS? The results of an international workshop on technology assessment of PACS. *Computer Methods & Programs in Biomedicine* 1992; 37(4): 265-71.

van Gennip EMSJ, Enning J, Fischer F, Glaser KH, Kilbinger M, Klose KJ, et al. Guidelines for cost-effective implementation of picture archiving and communication systems an approach building on practical experiences in three European hospitals. *International Journal of Bio-Medical Computing* 1996; 43(3): 161-78.

van Gennip EMSJ, Talmon JL, Bakker AR. ATIM, accompanying measure on the assessment of information technology in medicine. *Computer Methods & Programs in Biomedicine* 1994; 45(1-2): 5-8.

van Gennip EMSJ, van Poppel BM, Bakker AR, Ottes FP, Kouwenberg JML, Wilmink JBM. An analysis of the costs of a hospital-wide picture archiving and communication system with the software package CAPACITY. *European Journal of Radiology* 1991; 12(1): 69-78.

Warburton RN. Digital imaging at a community hospital: implications for hospital stays and teleradiology. *International Journal of Bio-Medical Computing* 1991; 28(3): 169-80.

Warburton RN, Fisher PD, Nosil J, Brauer GW, Lawrence WJ, Ritchie GW. Digital diagnostic imaging with a comprehensive PACS: hypothetical economic evaluation at a large community hospital. *Journal of Digital Imaging* 1990; 3(2): 101-7.

Weller G, Bluemke D, Gitlin J, Reichle R, Scott W, Kelen G, et al. A comparison of film amd screen intrepretations of radiographic examinations performed in the emergency department. In: Boehme J, Rowberg A, Wolfman N, eds. *SCAR 94 Computer Applications to Assist Radiology.* Boston: Symposia Foundation; 1994.

Williams SC, Contreras M, McBiles M, Cawthon MA, Shah RB. The impact of a picture archiving and communication system on nuclear medicine examination interpretation. *Journal Of Digital Imaging* 1997; 10(2): 51-6.

Wilson AJ, Mann FA, West OC, McEnery KW, Murphy WA, Jr. Evaluation of the injured cervical spine: comparison of conventional and storage phosphor radiography with a hybrid cassette. *Radiology* 1994; 193(2): 419-22.

Yamamoto LG, Inaba AS, Dimauro R. Personal computer teleradiology interhospital image transmission to facilitate tertiary pediatric telephone consultation and patient transfer: soft-tissue lateral neck and elbow radiographs. *Pediatric Emergency Care* 1994; 10(5): 273-7.

Appendix 1

Systematic Search and Appraisal Methods

METHODS

Systematic Reviews

The MDRC performed a systematic review of the published literature to address four significant issues related to PACS used in clinical settings: diagnostic accuracy, impact on clinical care, impact on organizational performance, and cost.

A *systematic* review differs from a traditional narrative literature review in that it uses a rigorous scientific approach to limit bias and to improve the accuracy of conclusions based on the available data (Guyatt, 1995). A systematic review addresses a focused clinical question, uses appropriate and explicit criteria to select studies for inclusion, conducts a comprehensive search, and appraises the validity of the individual studies in a reproducible manner.

Consistent with established methods for conducting a systematic review, the MDRC developed criteria to select studies for inclusion, conducted a comprehensive search, and appraised the validity of the individual studies in a reproducible fashion using the analytical frameworks presented below. Because preliminary bibliographic searches identified neither randomized clinical trials of PACS nor high-quality systematic overviews of PACS, the selection criteria were designed to include the best available empirical studies that addressed the relevant issues.

A. Inclusion and exclusion criteria

All published studies included in this report met the following inclusion criteria:

- English language journal articles reporting primary data obtained in clinical settings;
- study design and methods clearly described, with sufficient information to permit reproducibility;
- studies not duplicated or superseded by subsequent publications (with the same purpose) from the same research group;
- publication date of 1990 or later, in order to reflect the capabilities of this rapidly-evolving technology (Glass, 1992).

Additional inclusion criteria for studies of diagnostic accuracy:

- studies which meet the full or modified evidence-based medicine criteria (Table A1.1).

Additional inclusion criteria for studies of cost:

- studies which meet the HERG criteria for Stage II economic evaluation (Table A1.5).

B. Search strategy

Relevant literature was identified using formal search strategies. First, a broad search was conducted using the following terms: PACS, TELERADIOL$ (truncated to retrieve all variations), TELEMEDICINE AND RADIOLOG$, and RADIOLOGY INFORMATION SYSTEMS (the MeSH heading which includes PACS). The search was then narrowed to

those references covering: costs, benefits, trends, future, evaluation, innovation, organization and administration, assessment, forecasting, utilization, standards, statistics, numerical data, economics, supply, manpower, legislation, design, human factors, or human-computer. Searches were run on the National Library of Medicine's MEDLINE© and HEALTH Planning and Administration databases as well as on Elsevier publishing company's database, EMBASE© (1975 to June, 1997). To retrieve citations that would not yet have been entered into these databases, searches were also performed in all sections of Current Contents© Institute for Scientific Information (1990 to June, 1997). End-references from relevant studies, and assessments produced by other organizations were also included in the bibliographic retrieval.

Primary data reported from studies of the diagnostic accuracy of digital image systems used for teleradiology were also reviewed and included in this report. Teleradiology equipment can be viewed as PACS subsystems. Teleradiology systems support the acquisition, transmission, and viewing of images, but lack a PAC's capacity to archive imaging information.

C. Methodologic standards for studies

The purpose of appraising the literature using clearly defined methodologic criteria is to ensure that studies are evaluated in a consistent, reproducible manner, and that studies included in the report conform to established scientific standards. Studies reviewed for possible inclusion in this report were classified according to the strength of the evidence they provided, and the strongest available evidence for each significant issue was summarized in the report. The strength of a study is based on the overall research design and on the quality of the implementation and analysis.

The methodologic standards, and the types of studies to which they were applied, are summarized below. The standards are discussed in greater detail in the MDRC report *Assessing Diagnostic Technologies* (Flynn, 1996) and in references presented in the Selected Bibliography.

1. Studies of diagnostic accuracy

It is essential to demonstrate that PACS are sufficiently accurate for use in a clinical setting. The best evidence to determine if digital imaging (displayed on a workstation and/or on film) is as accurate as conventional film is to use both digital and analog imaging techniques on a representative group of patients from an appropriate population at the same point in time. Analyses of findings from diagnostic test studies are designed to determine how accurately a test can discriminate between disease and non-disease.

Evidence-based criteria have been established to guide the design and interpretation of studies of diagnostic tests, based on known sources of bias (Table A1.1). These criteria were used as standards by which to assess studies for potential inclusion in this report. Studies which do not meet these criteria are generally considered insufficiently rigorous to provide evidence on which to base patient care decisions. However, such studies can provide suggestive information which may be used to guide further planning and research efforts.

Table A1.1: Evidence-based medicine criteria for studies of diagnostic tests

1. clearly identified comparison groups, one or more of which is free of the target disorder
2. either an objective diagnostic standard or a contemporary clinical diagnostic standard with demonstrably reproducible criteria for any subjectively interpreted component
3. interpretation of the test without knowledge of the diagnostic standard result
4. interpretation of the diagnostic standard without knowledge of the test result

Source: Hayes and Sackett, eds., 1995.

A somewhat more refined set of quality standards were used to score diagnostic accuracy studies based on both study design and study implementation. The criteria were originally developed by the American College of Physicians to evaluate MRI literature (Table A1.2).

Table A1.2: Methodologic quality rating for diagnostic accuracy studies

Grade	*Criteria*	*Quality*
A	**Studies with broad generalizability to a variety of patients and no significant flaws in their research methods** • adequate sample size to provide sufficient statistical power • patients should be similar to those encountered in practice (that is, not only including those with severe disease) and from a group whose clinical symptoms are completely described • diagnosis defined by an appropriate reference standard (gold standard) • diagnostic studies technically of a high quality and evaluated independent of the reference diagnosis • prospective study design	**high quality evidence**
B	**Studies with a narrower spectrum of generalizability, and with only a few flaws that are well described so that their impact on conclusions can be assessed** • adequate sample size to provide sufficient statistical power • more limited spectrum of patients, typically reflecting the referral bias of university centers with more severely ill patients • free of other flaws in methods that promote interaction between test result and disease determination • prospective study design	**good quality evidence**
C	**Studies with several flaws in research methods, small sample size, or incomplete reporting** • retrospective study	**weak evidence**
D	**Studies with multiple flaws in research methods or reports of opinion unsubstantiated by data** • no credible reference standard for diagnosis • test result and determination of final diagnosis not independent • source of patients could not be determined, or source of patients obviously influenced by the test result, producing a work-up bias	**non-contributory evidence**

Sources: Kent, 1994; Kent and Larson, 1992; Kent and Larson, 1988.

Table A2.4 summarizes how studies of diagnostic accuracy which were included in this report were evaluated based on these two sets of standards.

2. Studies of processes and outcomes of care

Although these studies examine some variables which are unfamiliar to medical researchers, they in fact employ the same study designs as traditional medical research. Table A1.3 lists study designs for evaluating health care interventions, and ranks the persuasiveness of their findings, ordered from those providing the weakest to those providing the strongest evidence of a relationship between the intervention and the outcome of interest. Recommendations about using a technology should be linked to the quality of the available evidence, with the strength of the recommendation dependent on the quality of the available evidence. These traditional scientific standards for study design and strength of evidence were applied to studies reviewed for inclusion in the report, and studies which provided the strongest available evidence in each area of interest were selected for inclusion.

Table A1.3: Continuum of study designs and their causal implications

Study Design	***Inference / Strength of Evidence***
I. Anecdote Clinical hunches Case history	Speculative
II. Time series Ecologic correlations Cross-sectional	Suggestive
III. Case-control	Moderately suggestive
IV. Before-after with controls Historical cohort	Highly suggestive
V. Prospective cohort	Moderately firm
VI. Randomized controlled trials (RCT) Community randomized trials Systematic reviews of RTCs	Firm

Source: Adapted from Ibrahim, 1985.

Types I through III are observational studies, not true experiments. Observational studies are subject to many forms of bias which can diminish the accuracy of their findings. Therefore, they do not provide very persuasive evidence linking interventions with the outcomes observed. They *can* be useful for helping generate ideas for further research. Types IV and V studies are considered quasi-experimental designs. They are commonly used in healthcare (often because it is not possible to conduct true experiments with patients) and provide stronger evidence than can be obtained from observational studies. Type VI studies are true experiments, and provide the most persuasive evidence for linking interventions with the outcomes observed.

While this methodologic hierarchy is widely understood within the scientific community, methods for assessing the quality of data collected in studies of work process efficiency are less standardized and less familiar. Methods for assessing the most common outcome of

interest, time spent performing a task, are derived from the field of industrial engineering. Warburton (1992) classified several common methods of measuring units of work time, all of which have been used in estimating changes in work process related to PACS. These data quality rankings were applied to studies included in this report.

Table A1.4: Defining and Rating of Time Measurement Techniques

Technique	*Methods*	*Data Quality*
Time and motion study	direct observation by trained observer using stopwatch	•most accurate •most costly and disruptive
Activity sampling	•identify discrete activities of interest •observe each person at specified times •record which activity of interest staffer is then engaged in	•highly accurate with sufficient observations •less costly and disruptive
Self-recording Self-reporting	•record completion of task when done, or •report time for activity to interviewer	•less accurate than independent measurement •even less costly and less disruptive
Analysis of administrative data	•are often time estimation models based on time and motion studies •completeness, timeliness, and contents of administrative data may vary	•often fairly accurate, particularly for tracking changes over time •least costly and least disruptive

Source: Warburton, 1992.

3. Cost studies

Analytic frameworks discussed above provide a method to develop summary analyses based on the best quality data available in the published literature, but they do not explicitly address costs. Currently, there are no formally accepted criteria for synthesizing economic data from primary costing studies (Rigby, 1996).

Economic evaluations often use *ad hoc* methodologies, mixed research study designs, incorporate data obtained from earlier studies, or use modeling. When the methods are clearly described, the underlying study designs can be identified and may be classified within the framework presented in Table A1.3. This classification can then be used to rank the quality of the evidence (data) used in the study. This then provides a mechanism to evaluate the strength of the evidence used in the economic evaluation, and hence the reliability of the evaluation.

The above system ranks the strength of the evidence according to the *source* of the data used in the economic analysis. Economic analyses are also classified according to the *type* of data included in the analysis. The type of data used varies along three dimensions: perspective, type of analysis, and scope of costs and benefits included. The type of data used in an economic analysis should reflect the goals of the study.

Perspective: The perspective of the study defines the viewpoint from which the analysis is performed (for example, that of the patient, the service, or the institution). It determines the types of costs, consequences, and outcomes to be analyzed, as well as the conclusions to be drawn (Luce, 1995).

Measures of costs and benefits: The types of costs included in an analysis are classified as direct, indirect, or intangible. *Direct costs* are usually associated with monetary transactions and represent costs incurred in providing the care (payments for supplies, equipment, salaries). *Indirect costs* represent the value of consequences that cannot be counted as direct cost (number of work days lost, loss of productivity, loss of income due to premature death resulting from a missed diagnosis). *Intangible costs* (pain, anxiety, caregiver burden) are the remaining elements of the burden of the illness or activity. Although potentially significant, intangible costs are difficult to measure and are often omitted from economic analyses.

The measure used to value benefits in an economic evaluation is traditionally used as a means of classifying the study type. Measurement of the consequences of an expenditure may be omitted from a study, may be calculated in dollars, or may be expressed in the natural units of the benefits measured.

Types of economic analysis: Cost identification, cost-benefit, and cost-effectiveness analyses are three of the most common types of traditional economic evaluations.

A *cost identification analysis* simply measures costs of a specific activity (such as the implementation of a technology) without considering the consequences. A *cost-benefit analysis (CBA)* compares the costs of an activity to the dollar value of resources saved or created. A *cost-effectiveness analysis (CEA)* also measures the costs of an activity, but compares the costs to the benefits derived measured in their natural units (such as dollars per years of life saved) rather than in dollars. For a cost-benefit analysis or cost-effectiveness analysis to be done, outcomes thought to be a consequence of the activity being studied must be identified, and effectiveness must be measured.

The use of cost-effectiveness analysis has been widely recommended in the literature, perhaps because it appears to offer a method to compare returns on investments across different types of programs. However, CEAs have been difficult to conduct in the health care system, in which leaders often allocate resources to health care technologies which are still evolving, and for which strong evidence of effectiveness is still lacking.

Stages of economic analysis: A useful analytic framework has been proposed that frames economic evaluations in a way that appears better suited to the evolving nature of many health care technologies. The Health Economics Research Group at the University of York classifies economic evaluations into four stages (Sculpher, 1995). The stages capture the dynamic nature of technologies, and relate the level of maturity of the technology to the type and strength of evidence (as classified above) to be used in the economic evaluation. These stages can be viewed as analogous to the four main phases of experimentation used within the pharmaceutical industry (Pocock, 1983).

This framework incorporates the belief that all health care technology assessment, including economic evaluation, should be iterative, and that the evaluations themselves should represent a cost-effective use of resources. The HERG analytic framework appears to be well-suited for use with an evolving technology such as PACS, and is applied to studies reviewed for inclusion in this report.

Appendix 2

Evidence Tables for Diagnostic Accuracy Studies

Table A2.1: *Summary of the literature*
Diagnostic accuracy of PACS: comparison of conventional film to digital workstation imaging

Notes:

- *Three studies in this table (Paakkala, Goldberg, and DeCorato) used a case series design; these studies included patients determined to be free of disease who served as internal controls. The remaining studies used a case-control design and randomized presentation of images to interpreters.*
- *Except where otherwise indicated, ROC (Receiver Operating Characteristic) curves were used to compare diagnostic accuracy.*
- *The diagnostic standard (gold standard) in studies included in this table was the confirmed diagnosis by biopsy, CT, or expert opinion based on conventional film interpretation, clinical history, other testing, or response to treatment. Since blinding was inconsistent, the studies do not fully meet evidence-based criteria for diagnostic test evaluation.*
- *Four studies in this table (Slasky, Cox, Elam, and Thaete) included in their research design a comparison to a third imaging modality, that of digital imaging printed to hard copy (film). Findings related to digital image hard copy are reported in subsequent tables.*
- *Two studies (Ackerman, 1993 and Scott, 1995) were performed at the same institution, but address different diagnostic problems.*

Study	***Patients/Methods***	***Results/Comments***
Slasky et al., 1990 *Departments of Diagnostic Radiology and Statistics, University of Pittsburgh*	**Purpose** to compare the diagnostic accuracy of interpreting adult chest images using three formats: conventional film images, digital image hard copy (film), and digital image soft copy (on high resolution monitors) **Cases/Controls** • case sample of 300 prospectively identified, high quality chest radiographs from an outpatient facility, 113 of which were normal • abnormal cases represented a range of severity of the following conditions: any abnormality, lung nodules, interstitial disease, and/ or pneumothorax **Methods** • conventional films were digitized with a high-contrast-sensitivity high-resolution digitizer for viewing on workstations and laser-printing to film • 7 board-certified radiologists interpreted each of 300 chest images twice: once on conventional film and once on workstation. 5 of the radiologists also interpreted digital hard copy (film) version of images • images presented in random order • readers blinded to patient clinical information • use of workstation tools to change window (contrast) and level (brightness) was encouraged • gold standard: consensus of 2 experts based on conventional film imaging, and independently corroborated by biopsy, CT, or follow-up exam • spatial resolution of digital images: 2048 x 2400 x 12 bit digital images displayed on 1536 x 2048 x 8 bit monitor or printed to high-resolution film	**Image Analysis** • overall accuracy in the detection of “any abnormality”, interstitial disease, and pneumothorax significantly higher for film than for digital image viewed on workstation. No significant difference between these two modalities for accuracy of detection of pulmonary nodules • time required to diagnose cases was comparable for all modalities **Authors’ Comments** • as a group, radiologists did not perform as well with either hard-copy or soft-copy of digital images as with conventional film when very high resolution was needed for disease detection • extensive training (not provided in this study) may be required with high-resolution workstations before radiologists perform at optimal level ***Additional study findings reported in other tables***

Study	Patients/Methods	Results/Comments
Cox et al., 1990 *University of Kansas Medical Center*	**Purpose** to compare the diagnostic accuracy of interpreting adult chest images using three formats: conventional film images, digital images printed as hard copy (film), and digital images displayed on high resolution monitors **Cases/Controls** • 163 chest radiographs selected for study: 99 which demonstrated one or more of 9 pulmonary abnormalities with a range of diagnostic difficulty and 64 normal chest images **Methods** • conventional film images were digitized for viewing on workstations or for laser printing to create hard copy (film) of digitized images • readers were trained to use workstations, scoring form, and defined diagnostic criteria for selected abnormalities • six board-certified radiologists, three with no prior computer or workstation experience, analyzed images • readers were blinded to other clinical information • each read 163 images in random order: 1/3 as conventional film, 1/3 as digital image hard copy, and 1/3 on workstation. No radiologist viewed more than one version of any case • both conventional and laser-printed films were viewed on a film alternator • zoom and pan functions available on workstations; use of image manipulation techniques while reading "soft copy" not quantified • gold standard: consensus of 2 experts, based on imaging • spatial resolution of digital images: 2048 x 2048 x 12 bit digital images displayed on 2048 x 2560 (high resolution) monitor or printed to film. Two versions of film were printed, one with higher-contrast; study coordinator selected the image most closely approximating the original film	**Image Analysis** • accuracy of conventional film and workstation images equivalent for five of the conditions of interest: costophrenic angle blunting, atelectasis, consolidation, apical scarring, and hilar/ mediastinal masses • conventional film more accurate than digital image soft copy (workstation) for detecting pneumothorax and interstitial disease • workstation images more accurate than conventional film for detecting parenchymal masses **Authors' Comments** • present technology, as tested, may not be adequate for primary diagnoses of interstitial pulmonary disease and pneumothorax • workstations may offer some advantage over film. Observers were able to detect parenchymal nodes more accurately on workstations than on conventional film • half of readers in this study had no prior experience with workstations. Increased familiarity with equipment, and recent improvements in workstations, are expected to further improve performance of readers on workstations ***Additional study findings reported in other tables***
Paakkala et al., 1991 *Tampere University Central Hospital, Finland*	**Purpose** to evaluate the image quality and clinical practicability of teleradiology for interpretation of chest and bone imaging **Cases/Internal Controls** • case series of 372 images, prospectively studied: - 146 chest (38% normal) - 226 bone (58% normal) **Methods** • conventional film images were digitized, transmitted over digital telephone line (64 Kbits/second), and viewed on a monitor • readers were not blinded to clinical information on X-ray request forms • a radiologist and a third-year resident interpreted digital images, and evaluated image quality. Radiologist read 190 images, resident read all images. 3 months later, same physicians read original films • gold standard: final diagnosis made from review of images and medical records • difference in accuracy between modalities calculated mathematically (without ROC curves) • spatial resolution of digital images: 512 x 512 pixels x 8 bits	**Image Analysis** *Note that authors do not indicate level of statistical significance of their findings, nor are operating characteristics reported by specific diagnosis* • 90% of films and 73% of digitized images viewed on monitors were considered satisfactory by reader(s) • overall, the interpretation of film images was 5% more sensitive and 3% more specific than interpretation of monitor images • for chest films, the accuracy of interpretation was 2.2% higher than for monitor images; for bone films, accuracy of interpretation was 4.7% higher than for monitor images • zooming, brightness or contrast control judged to be useful in 21% of cases **Authors' Comments** • quality of digital image transmission improved as operators gained experience with equipment • no major discrepancies between the two imaging modalities arose in 96% of cases, and, despite the slightly poorer accuracy of transmitted images, it was an acceptable means of augmenting the usefulness of the radiologist • senior consultant used zooming, brightness control, and contrast adjustment more than resident

Study	Patients/Methods	Results/Comments
Elam et al., 1992 *University of Arizona College of Medicine; University of California, San Francisco*	**Purpose** to compare the diagnostic accuracy of interpreting pneumothorax on adult chest images using four formats: conventional film images, digital images printed on film in conventional small format (17.8 x 21.6 cm) or large format (35.6 x 43.1 cm), and digital images displayed on a monitor **Cases/Controls** • 45 chest radiographs were selected - 23 demonstrating a spectrum of pneumothoraces - 22 age-matched controls with either normal chests or abnormality other than pneumothorax **Methods** • matched film-screen (analog) and computed radiography (digital) images were obtained from 45 patients for viewing in four formats • 5 board-certified radiologists, all experienced with digital images and trained in image manipulation techniques, interpreted all images in each of 4 formats • images presented in random order. Images on monitors were read in one session; film images were presented in sets of 45 paired images, with 1-2 weeks given for interpretation of each set • window, level, and zoom functions available on workstations, but their use was not quantified • readers indicated presence or absence of each pulmonary abnormalities listed for them, and rated degree of confidence in their diagnosis • readers blinded to clinical diagnosis • gold standard: diagnosis as determined by follow-up imaging and by patient's clinical course • spatial resolution of digital images: 1760 x 2140 pixel resolution of initial digital image, printed without loss or compression to film or displayed on a 1024 x 1536 pixel monitor	**Image Analysis** • no significant difference in accuracy, sensitivity or specificity of detection of pneumothorax between conventional film and digital images interpreted on workstations • the differences in sensitivity between film (.82) and workstations (.65) approached significance (p=.06) **Authors' Comments** • based on author's post-hoc calculations, study did not have the power to detect significant differences in accuracy if, in fact, they existed • the small study size, or greater experience of the study's readers with digital chest radiography, may explain difference in outcome between this and the Slasky study ***Additional study findings reported in other tables***
Franken et al., 1992 *Department of Radiology, University of Iowa*	**Purpose** to compare the diagnostic accuracy of interpreting clinical neonatal radiographs using conventional film images vs. digital workstation **Cases/Controls** • case sample of 100 chest or abdominal films from neonatal ICU, representing a wide range of subtle manifestations of common pediatric diseases - 58 chest images: 14 normals, 4-10 images of each of six pulmonary diseases - 42 abdominal images: 19 normals, 2-16 images of each of four abdominal diseases **Methods** • film images were digitized for viewing on a PACS workstation • original conventional x-ray films were viewed with a hot light as well as a view box • 4 radiologists rated images as normal or abnormal, and ranked degree of confidence in their conclusions. Each radiologist read all images in both film and workstation format • images were presented in random order • 2 radiologists blinded to patient age, 2 aware of patient age • gold standard used: confirmed diagnosis based on demonstrated typical clinical course of disease; subsequent imaging reflecting fulminating disease; or alternative diagnostic testing with a higher degree of accuracy • spatial resolution of workstation: 1024 x 1024 pixels. Small pediatric images could be displayed at full resolution, with imaging comparable to that of a 2000 x 2000 pixel monitor	**Image Analysis** • "no appreciable difference" in accuracy of interpretation between conventional film and digitized image viewed on a PACS workstation **Authors' Comments** • although no differences between film and workstations were found, the confidence intervals for the differences were too large to conclude that the two were truly comparable • smaller size of pediatric image allowed complete images to be displayed at full resolution on PACS workstations. To produce comparable spatial resolution, 2048 x 2048 pixel monitors may be required for displaying digital images of adult chest films

Study	Patients/Methods	Results/Comments
Ackerman et al., 1993 *Johns Hopkins Medical Institution*	**Purpose** to evaluate the diagnostic accuracy of conventional film as compared with digitized images displayed on a workstation for detecting fractures and pneumonia **Cases/Controls** 160 films selected: 40 fracture cases and 40 pneumonia cases with 80 age- and gender-matched controls **Methods** • selected conventional film images were digitized and transmitted via a T1 telephone line (1.544 Mbits/second) for viewing on a remote workstation • readers received training in use of workstations, including use of image manipulation tools (including zoom and brightness-contrast) • 8 radiologists (4 faculty, 2 fellows, and 2 senior residents) read images, rated degree of confidence in their diagnosis, and assessed image quality • each reader viewed half the examinations as film and half as digital images; no reader interpreted the same image from both film and digital mode. No reader was blinded to clinical or demographic information • gold standard: diagnosis based on interpretation of conventional film images, determined by consensus panel of 6 experts. Level of diagnostic difficulty and technical quality of images were rated by same panel • spatial resolution of monitor: 1280 x 1024 pixels	**Image Analysis** • overall sensitivity significantly higher for conventional film than for workstation images (89% vs. 78%) • overall specificity of the two modes (film vs. workstation) not significantly different • overall accuracy of interpretations significantly higher for conventional film than for workstation • accuracy of interpretations significantly increased with reader's increased confidence in diagnosis, as well as with increased quality of image (92% vs. 85%) • no significant difference in overall performance between radiology faculty and fellows/ residents • no increase in the accuracy of interpreting workstation images over time **Authors' Comments** • for the selected conditions, this workstation configuration is inappropriate for primary diagnostic interpretation • readers emphasized the need for faster workstations with higher spatial resolution • based on post-hoc analysis, the size of the subgroups in the study was too small, and the analysis may have failed to detect some of the differences between conventional film and workstation images
Goldberg et al., 1993 *Massachusetts General Hospital and Harvard Medical School*	**Purpose** to compare the accuracy of interpretation of conventional film, as compared with digitized images displayed on high-resolution workstations **Cases/Internal Controls** • case series of the first 685 plain films transmitted from an outpatient center to the hospital radiology department via high-speed teleradiology system: - 530 adults: 334 normal, 196 abnormal chest, bone and soft tissue, sinus, or KUB - 205 pediatric: 119 normal, 86 abnormal chest, bone and soft tissue, or sinus **Methods** • conventional film images were digitized and transmitted via T1 telephone lines to workstation accompanied by: previous images (when available), clinical data, and demographic data • hospital radiologists interpreted the images on workstations and faxed preliminary reports to the outpatient site • original film images "usually" interpreted by different radiologist • readers able to access workstation interpretations when making diagnosis based on film • gold standard: diagnosis based on conventional film reading, with arbitration panel review of discrepant cases. Preliminary diagnoses based on workstation image interpretation available to readers who interpreted conventional film images • spatial resolution of images: 1684 x 2048 x 12 bit digital images displayed on a 2048 x 2560 x 8 bit monitor	**Image Analysis** • overall interpretive accuracy of use of digital workstation compared with that of film was 98% • overall diagnostic sensitivity was 96% and overall specificity was 99% • overall accuracy of interpretations significantly higher for film than for workstation • review of discrepant cases identified 14 errors attributed to workstation interpretation and 3 errors in conventional film interpretation **Authors' Comments** • accurate primary diagnosis with high-resolution digital teleradiology is now feasible • a higher resolution system than that tested here might be needed for some diagnostic tasks

Study	Patients/Methods	Results/Comments
Thaete et al., 1994 *Department of Radiology, University of Pittsburgh*	**Purpose** to compare the diagnostic accuracy of interpreting adult chest images using three formats: conventional film images, digital images printed as hard copy (film), and digital images displayed on a workstation **Cases/Controls** • 310 chest radiographs prospectively selected from outpatient facility: - 198 had one or more abnormality (24 alveolar infiltrate, 124 interstitial disease, 84 nodules, 27 pneumothorax, 18 rib fractures) - 112 controls had none of abnormalities of interest **Methods** • matched film-screen (analog) and high-resolution-high-contrast computed radiography (digitized) images were obtained from same patients • readers trained in use of workstations, study forms, and use of defined diagnostic criteria for selected abnormalities • 9 board-certified radiologists analyzed images. 7 had extensive experience with workstations and 2 were trained prior to the study • readers were blinded to other clinical information • each reader interpreted 30-40 images per session. Order of images, and type of format used per session (workstation, conventional film or digitized image on film), was randomized. All radiologists read all films in all formats • gold standard: consensus of 2 experts based on biopsy, CT or other follow-up imaging, and clinical history • spatial resolution of digital images: 4096 x 5000 x 12 bit digital images displayed on 1536 x 2048 x 8 bit monitor or printed to film with high-resolution printer	**Image Analysis** • overall accuracy in the detection of alveolar infiltrate, interstitial disease and pneumothorax significantly higher for conventional film than for digital image displayed on workstation • no significant difference between these two modalities for detection of nodules • no significant difference between these two modalities for detection rib fractures *it should be noted that study lacked the power to detect such a difference if, in fact, it existed* **Authors' Comments** • although digital imaging of the chest has improved, conventional film imaging technology has also improved • as a group, radiologists did not perform as well with digitally acquired images displayed on workstations as with film for conditions such as interstitial disease and pneumothorax. • workstations used may not be adequate for primary diagnosis of these abnormalities at this time ***Additional study findings reported in other tables***
Korsoff et al., 1995 *Turku University Central Hospital and Technical Research Center of Finland*	**Purpose** to evaluate the diagnostic reliability of transmitted digitized images compared with original chest films **Cases/Controls** • 43 cases: 15 with subtle pneumothorax, 12 with interstitial lung disease, and 16 controls • method of control selection not described **Methods** • prospectively identified, selected conventional film images were digitized using both a 1024 x 1024 x 12 bit and a 2048 x 2048 x 12 bit matrix, and transmitted by modem (64 Kbits/second) to workstations • 5 hospital radiologists interpreted the images and rated degree of confidence in reading according to the following schedule: month 1: 1024 x 1024 pixel images; month 2: 2048 x 2048 pixel images; month 3: original films • all radiologist read all images in all formats • readers were not fully blinded. All images were accompanied by the same clinical information: "suspected pneumothorax or interstitial pulmonary disease" • gray-scale windowing and zooming functions available on monitors • all 2K x 2K images viewed using zoom function • gold standard: consensus of 3 radiologists based on reading original films and confirmed with strong clinical and imaging follow-up • spatial resolution of monitor used to display all digital images: 1280 x 1024 x 8 bits	**Image Analysis** • no significant differences in the detection of interstitial fibrosis between conventional films and digital images of either moderate or high resolution • no significant differences found in the detection of pneumothorax between conventional films and high resolution (2048 x 2048) digital images displayed on the 1280 x 1024 x 8 bit monitor using magnification • detection of pneumothorax significantly higher for conventional film than for moderate resolution (1024 x 1024) digital images displayed on the 1280 x 1024 x 8 bit monitor • gray-scale windowing not used by readers because of function's slow response time **Authors' Comment** • accurate primary diagnosis of pulmonary disease is possible with 2048 x 2048 pixel matrix, and 1024 x 1024 matrix may be adequate for use by an experienced radiologist

Study	Patients/Methods	Results/Comments
Scott et al., 1995 *Johns Hopkins Medical Institution*	**Purpose** to compare the accuracy of interpretation, by radiologists and ER physicians, of conventional film and digital images displayed on workstations **Cases/Controls** • 120 films retrospectively selected: - 60 diagnostically difficult cases of musculoskeletal, abdominal, and chest disorders - 60 controls • frequency distribution of disorders similar to that seen in ER **Methods** • selected film images were digitized for viewing on a workstation • contrast, brightness, and zoom control available on workstation; use not quantified • original conventional x-ray films were read on a view box • readers: 4 faculty radiologists, 4 second-year radiology residents, 4 faculty ER physicians, 4 second-year ER residents • each reader interpreted 60 digital images and 60 film images of different cases; each case was viewed only once by a reader • images were presented in random order • readers not blinded to some clinical information • gold standard: diagnosis based on conventional film image. Discrepant interpretations were reviewed by expert panel • spatial resolution of digital images: 1200 x 1600 pixels	**Image Analysis** • overall accuracy significantly higher for conventional film than for workstation images (64.5% vs. 57.3%). This difference in accuracy varied with specialty area of the reader and with difficulty of the diagnostic task • overall sensitivity significantly higher for conventional film than for workstation images (50.2% vs. 39.6%) • no significant difference in overall specificity between the two imaging modalities • when data were analyzed by specialty area: - for ER physicians, no significant difference in accuracy of reading from film vs. workstations - for radiologists, accuracy significantly higher for film than for workstation images - radiologists' interpretation of workstation images more accurate than ER physicians' interpretation of film • when data analyzed by diagnostic difficulty: - for cases of low difficulty, reader performance was relatively high and essentially equal for film and workstation interpretation - for cases of moderate difficulty, interpretation of film images was significantly more accurate than for workstation interpretation (61.7% vs. 49.4%) and significantly more sensitive (54.7% vs. 40.5%) - for cases of high difficulty, interpretation of film images was more accurate, sensitive, and specific than for workstation image interpretation, but differences were not statistically significant **Authors' Comments** • the equipment configuration tested was not acceptable for making primary diagnoses of ER exams; high spatial resolution is critical for workstation imaging of subtle abnormalities • the extensive clinical information normally used by ER physicians when interpreting images was not provided to readers in this study

Study	Patients/Methods	Results/Comments
DeCorato et al., 1995 *St. Luke's-Roosevelt Hospital Center, New York*	**Purpose** to assess the efficacy of a digital teleradiology system in the interpretation of ER radiology studies **Cases/Internal Controls** • prospectively assembled case series of 812 imaging examinations obtained over 6 months on night shift in ER, 46% of which were abnormal, 24 of which represented subtle abnormalities • included 693 conventional film, 118 CT, one MRI **Methods** • all images were printed to film, then digitized. Image data was transmitted from one hospital ER to a remote hospital via a T1 telephone line • readers were radiology residents with 1-3.5 years experience, all with prior training and experience in interpreting workstation images • image manipulation functions available on workstations; use not quantified • gold standard: official film interpretation performed within 24 hours by a board-certified radiologist who had not read workstation images • blinding of readers of workstation images to clinical information not reported • discrepancies between workstation image interpretations by residents and film image interpretation by radiologists evaluated for potential clinical significance with assistance of ER physician • discrepant cases of clinical significance re-digitized and evaluated by experts blinded to diagnoses • data from conventional film images which were digitized for transmission not analyzed separately from MRI/CT data. No ROC analysis performed • spatial resolution of images: 2040 x 2056 x 12 bit images displayed on a 1200 x 1600 pixel monitor	**Image Analysis** *Note that authors do not indicate the level of statistical significance of their findings* • 95% (95% CI=93.6%-96.6%) of transmitted cases had no clinically significant differences in interpretation between workstation image interpretation by residents and film interpretation by radiologists • half of discrepant readings: - 19 cases were judged to be reader error by residents interpreting digital images - 14 were judged to reflect interobserver variability - 3 were considered to reflect inadequate digital imaging - 2 were judged to be film image interpretation error **Authors' Comments** • tested system "can be both reliable and effectively used" in off-hour interpretation of ER studies • image manipulation on workstations was judged to be slow and cumbersome • findings consistent with those of Goldberg

Table A2.2: *Summary of the literature*
Diagnostic accuracy of PACS: comparison of conventional film and digital images as hard copy (film)

Notes:

- *Studies in this table used case-control design and randomized presentation of images to interpreters.*
- *ROC (Receiver Operating Characteristic) curves were used to compare diagnostic accuracy.*
- *The diagnostic standard (gold standard) used was confirmed diagnosis by biopsy, CT, or expert opinion based on review of conventional analog images and/ or medical records.*
- *Studies included in this table met most or all of the evidenced-based criteria for diagnostic test evaluation; studies by Elam, Yoshino, and Kondoh do not explicitly indicate that readers were blinded.*
- *Two studies (Slasky and Thaete) were performed by different departments in the same institution, using different subjects and methods.*

Study	***Patients/Methods***	***Results/Comments***
Slasky et al., 1990 *Departments of Diagnostic Radiology and Statistics, University of Pittsburgh*	**Purpose** to compare the diagnostic accuracy of interpreting adult chest images using three formats: conventional film images, digital image hard copy (film), and digital image soft copy (on high resolution monitors) **Cases/Controls** • case sample of 300 prospectively identified, high quality chest radiographs from an outpatient facility, 113 of which were normal • abnormal cases represented a range of severity of the following conditions: any abnormality, lung nodules, interstitial disease and/ or pneumothorax **Methods** • conventional films were digitized with a high-contrast-sensitivity high-resolution digitizer for viewing on workstations and laser-printing to film • 7 board-certified radiologists interpreted each of 300 chest images twice: once on conventional film and once on workstation. 5 of the radiologists also interpreted digital hard copy (film) version of images • images presented in random order • readers blinded to patient clinical information • use of workstation tools to change window (contrast) and level (brightness) was encouraged • gold standard: consensus of 2 experts based on conventional film imaging, and independently corroborated by biopsy, CT, or follow-up exam • spatial resolution of digital images: 2048 x 2400 x 12 bit digital images displayed on 1536 x 2048 x 8 bit monitor or printed to high-resolution film	**Image Analysis** • overall accuracy in the detection of pneumothorax significantly higher for conventional film than for hard copy (film) of digital image. No significant difference between these two modalities for accuracy of detection of “any abnormality”, interstitial disease, or pulmonary nodules **Authors’ Comment** • as a group, radiologists did not perform as well with hard-copy of digital images as with conventional film when very high resolution was needed for disease detection ***Additional study findings reported in other tables***

Study	Patients/Methods	Results/Comments
Cox et al., 1990 *University of Kansas Medical Center*	**Purpose** to compare the diagnostic accuracy of interpreting adult chest images using three formats: conventional film images, digital images printed as hard copy (film), and digital images displayed on high resolution monitors **Cases/Controls** • 163 chest radiographs selected for study: 99 which demonstrated one or more of 9 pulmonary abnormalities with a range of diagnostic difficulty and 64 normal chest images **Methods** • conventional film images were digitized for viewing on workstations or for laser printing to create hard copy (film) of digitized images • readers were trained to use workstations, scoring form, and defined diagnostic criteria for selected abnormalities • 6 board-certified radiologists, three with no prior computer or workstation experience, analyzed images • readers were blinded to other clinical information • each read 163 images in random order: 1/3 as conventional film, 1/3 as digital image hard copy, and 1/3 on workstation. No radiologist viewed more than one version of any case • both conventional and laser-printed films were viewed on a film alternator • zoom and pan functions available on workstations; use of image manipulation techniques while reading "soft copy" not quantified • gold standard: consensus of 2 experts, based on conventional imaging • spatial resolution of digital images: 2048 x 2048 x 12 bit digital images displayed on 2048 x 2560 (high resolution) monitor or printed to film. Two versions of film were printed, one with higher-contrast; study coordinator selected the image most closely approximating the original film	**Image Analysis** • no significant difference between conventional film and hard copy (film) of digital image for the detection of 8 of the 9 abnormalities tested (costophrenic angle blunting, atelectasis, consolidation, apical scarring, hilar/mediastinal mass, obstructive disease, pneumothorax, and interstitial disease) • digital image hard copy more accurate than conventional film for detecting parenchymal masses **Authors' Comment** • performance of 2048 x 2048 hard copy of digital imaging is generally equivalent to that of conventional radiography, with the exception of the detection of parenchymal masses. The improved rendition of contrast information in the digitized hard copy may contribute to this finding ***Additional study findings reported in other tables***

Study	Patients/Methods	Results/Comments
Elam et al., 1992 *University of Arizona College of Medicine; University of California, San Francisco*	**Purpose** to compare the diagnostic accuracy of interpreting pneumothorax on adult chest images using four formats: conventional film images, digital images printed on film in conventional small format (17.8 x 21.6 cm) or large format (35.6 x 43.1 cm), and digital images displayed on a monitor **Cases/Controls** • 45 chest radiographs were selected: - 23 demonstrating a spectrum of pneumothoraces - 22 age-matched controls with either normal chests or abnormality other than pneumothorax **Methods** • matched film-screen (analog) and computed radiography (digital) images were obtained from 45 patients for viewing in four formats • 5 board-certified radiologists, all experienced with digital images and trained in image manipulation techniques, interpreted all images in each of 4 formats • images presented in random order. Images on monitors were read in one session; film images were presented in sets of 45 paired images, with 1-2 weeks given for interpretation of each set • window, level, and zoom functions available on workstations, but their use was not quantified • readers indicated presence or absence of each pulmonary abnormalities listed for them, and rated degree of confidence in their diagnosis • readers blinded to clinical diagnosis • gold standard: diagnosis as determined by follow-up imaging and by patient's clinical course • spatial resolution of digital images: 1760 x 2140 pixel resolution of initial digital image, printed without loss or compression to film or displayed on a 1024 x 1536 pixel monitor	**Image Analysis** • no significant difference in accuracy, sensitivity, or specificity of detection of pneumothorax between conventional film and digital images printed to either small or large format film **Authors' Comments** • based on author's post-hoc calculations, study did not have the power to detect significant differences in accuracy if, in fact, they existed • the small study size, or greater experience of the study's readers with digital chest radiography, may explain difference in outcome between this and the Slasky study ***Additional study findings reported in other tables***
Yoshino et al., 1992 *Arizona Health Sciences Center, University of Arizona, and Boston Children's Hospital*	**Purpose** to assess the diagnostic accuracy of high-resolution teleradiology for detecting cervical spine fractures **Cases/Controls** •50 films selected retrospectively for study - 25 patients with cervical spine fractures of varying diagnostic difficulty - 25 controls (5 normal, 20 with degenerative disc disease) **Methods** • conventional film images were digitized, transmitted to a remote site via a T1 telephone line (1.54 megabytes/second), and laser-printed to film to produce hard-copy of the digital image • all images used in the study were judged to be of acceptable quality • digital film size was 56% of conventional film image size • each image was examined by 4 readers: 2 neuroradiologists, a general radiologist, and a radiology fellow. Each reader interpreted all images, with at least 4 weeks between reading of original and printed version of transmitted image • readers evaluated film for presence or absence of fracture and rated their degree of certainty in diagnosis • blinding of readers to clinical information not indicated • gold standard used was one of the following: visualization of bone at surgery or autopsy, CT, or follow-up X-ray indicating healing fracture • spatial resolution of digital images: 2048 x 2048 x 8 bits	**Image Analysis** • for the 2 experienced readers, conventional film significantly more accurate than 56%-sized hard copy of digital images in the detection of cervical neck fractures • for the 2 less-experienced readers, no significant difference in accuracy between conventional film and reduced-sized digital film • the diagnostic performance of neuroradiologists using transmitted images was better than that of general radiologists using conventional imaging, although the difference did not reach statistical significance **Authors' Comments** • high resolution alone may not be adequate to give teleradiology images equal diagnostic performance to original film images for cervical spine fracture detection. It may be necessary to view transmitted images on workstations rather than on film in order to have the capacity to manipulate images on a monitor to improve diagnostic capabilities of the system • the decision about whether or not to use teleradiology needs to be made in light of both the system itself and the users of the system

Study	Patients/Methods	Results/Comments
Thaete et al., 1994 *Department of Radiology, University of Pittsburgh*	**Purpose** to compare the diagnostic accuracy of interpreting adult chest images using three formats: conventional film images, digital images printed as hard copy (film), and digital images displayed on a workstation **Cases/Controls** • 310 chest radiographs prospectively selected from outpatient facility: - 198 had one or more abnormality (24 alveolar infiltrate, 124 interstitial disease, 84 nodules, 27 pneumothorax, 18 rib fractures) - 112 controls had none of abnormalities of interest **Methods** • matched film-screen (analog) and high-resolution-high-contrast computed radiography (digitized) images were obtained from same patients • readers trained in use of workstations, study forms, and use of defined diagnostic criteria for selected abnormalities • 9 board-certified radiologists analyzed images. 7 had extensive experience with workstations and 2 were trained prior to the study • readers were blinded to other clinical information • each reader interpreted 30-40 images per session. Order of images, and type of format used per session (workstation, conventional film, or digitized image on film), was randomized. All radiologists read all films in all formats • gold standard: consensus of 2 experts based on biopsy, CT or other follow-up imaging, and clinical history • spatial resolution of digital images: 4096 x 5000 x 12 bit digital images displayed on 1536 x 2048 x 8 bit monitor or printed to film with high-resolution printer	**Image Analysis** • overall accuracy in the detection of interstitial disease significantly higher for conventional film than for hard-copy (film) of digital image. No significant difference between these two modalities for detection of alveolar infiltrate, nodules, pneumothorax, or rib fracture **Authors' Comments** • although digital imaging of the chest has improved, conventional film imaging technology has also improved • as a group, radiologists did not perform as well with digitally acquired images displayed on workstations as with film or hard-copy of digital images for conditions such as interstitial disease and pneumothorax. Workstations used may not be adequate for primary diagnosis of these abnormalities at this time ***Additional study findings reported in other tables***
Kondoh et al., 1994 *Osaka University*	**Purpose** to compare the diagnostic accuracy of interpreting chest images using conventional film vs. two formats of hard copy (film) of digital images: full and 2/3 size **Cases/Controls** • 40 chest radiographs selected for study: - 20 with a variety of subtle interstitial pulmonary lesions (N for individual disorders ranged from 1-8) - 20 control patients without pulmonary lesions • method of image selection not described **Methods** • digital images were captured via computed radiography (CR) with a 2000 x 2000 x 10-bit format • 11 radiologists (7 experienced, 4 junior staff or residents) read all cases in each of three formats (conventional film and 2 sizes of CR film) • readers identified abnormalities and rated degree of confidence in their diagnosis • readers not blinded to types or severity of abnormalities included in study • blinding of CT readers to diagnostic test information not indicated • order and timing of image presentation not described • gold standard: computed tomography (CT) of the chest • spatial resolution of digital images: 2000 x 2000 pixels	**Image Analysis** • conventional film significantly more accurate than full-sized hard copy (film) printed from digital images • for all readers combined, no significant difference in accuracy of interpretation between full-sized and 2/3-sized hard copy of digital images • for all readers combined, no significant difference in accuracy of interpretation between conventional film and 2/3-sized hard copy of digital images • when data from experienced and inexperienced radiologists were analyzed separately: - for inexperienced radiologists, no significant difference in accuracy of interpretation among 3 types of images (conventional film, full-sized, and 2/3-sized hard copy of digital images) - for experienced radiologists, interpretation from conventional film significantly more accurate than from either sized hard copy of digital image **Authors' Comments** • based on subgroup analysis, the image quality needed to detect subtle abnormalities was dependent on level of reader's experience • early detection of acute subtle interstitial lung abnormalities is important in the clinical setting in which this study was performed. The digital imaging system tested was not fully satisfactory for the primary diagnosis of subtle pulmonary disease • higher resolution imaging equipment was purchased to better meet the needs of these patients ***Additional study findings reported in other tables***

Table A2.3: *Summary of the literature*
Diagnostic accuracy of PACS: comparison of digital images viewed as hard copy (on film) to digital images viewed as digital images viewed as soft copy (on workstation)

Notes:

- *The studies included in this table used a case-control design and randomized presentation of images to the readers.*
- *ROC (Receiver Operating Characteristic) curves were used to compare diagnostic accuracy.*
- *The diagnostic standard (gold standard) used was consensus of experts based on all clinical and imaging information, or (for Straub et al.) surgical confirmation when available.*
- *Since blinding of readers is not specifically addressed in the methods sections, it is unclear if the studies included in this table fully meet evidence-based criteria for diagnostic test evaluation.*

Study	***Patients/Methods***	***Results/Comments***
Cox et al., 1990 *University of Kansas Medical Center*	**Purpose** to compare the diagnostic accuracy of interpreting adult chest images using three formats: conventional film images, digital images printed as hard copy (film), and digital images displayed on high resolution monitors **Cases/Controls** • 163 chest radiographs selected for study: 99 which demonstrated one or more of 9 pulmonary abnormalities with a range of diagnostic difficulty, and 64 normal chest images **Methods** • conventional film images were digitized for viewing on workstations or for laser printing to create hard copy (film) of digitized images • readers were trained to use workstations, scoring form, and defined diagnostic criteria for selected abnormalities • 6 board-certified radiologists, 3 with no prior computer or workstation experience, analyzed images • readers were blinded to other clinical information • each read 163 images in random order: 1/3 as conventional film, 1/3 as digital image hard copy, and 1/3 on workstation. No radiologist viewed more than one version of any case • both conventional and laser-printed films were viewed on a film alternator • zoom and pan functions available on workstations; use of image manipulation techniques while reading "soft copy" not quantified • gold standard: consensus of 2 experts, based on imaging • spatial resolution of digital images: 2048 x 2048 x 12 bit digital images displayed on 2560 x 2048 (high resolution) monitor or printed to film. Two versions of film were printed, one with higher-contrast; study coordinator selected the image most closely approximating original film	**Image Analysis** • no significant difference between digital hard copy (film) and soft copy (workstation) for the detection of costophrenic angle blunting, atelectases, consolidation, apical scarring, and hilar/mediastinal/parenchymal masses • digital image hard copy (film) more accurate than digital image soft copy (workstation) for detecting obstructive airway disease, interstitial disease, and pneumothorax **Authors' Comments** • present technology, as tested, may not be adequate for primary diagnoses of interstitial pulmonary disease and pneumothorax • half of the readers in this study had no prior experience with workstations. Increased familiarity with equipment, and recent improvements in workstations, are expected to further improve performance of readers on workstations ***Additional study findings reported in other tables***

Study	Patients/Methods	Results/Comments
Elam et al., 1992 *University of Arizona College of Medicine; University of California, San Francisco; University of Virginia; Kansas Medical Center*	**Purpose** to compare the diagnostic accuracy of interpreting pneumothorax on adult chest images using four formats: conventional film images, digital images printed on film in conventional small format (17.8 x 21.6 cm) or large format (35.6 x 43.1 cm), and digital images displayed on a monitor **Cases/Controls** • 45 chest radiographs were selected: - 23 demonstrating a spectrum of pneumothoraces - 22 age-matched controls with either normal chests or abnormality other than pneumothorax **Methods** • matched film-screen (analog) and computed radiography (digital) images were obtained from 45 patients for viewing in four formats • 5 board-certified radiologists, all experienced with digital images and trained in image manipulation techniques, interpreted all images in each of 4 formats • images were presented in random order. Images on monitors were read in one session; film images were read in sets of 45 paired images, with 1-2 weeks for reading each set • window, level, and zoom functions available on workstations, but their use was not quantified • readers indicated presence or absence of each pulmonary abnormalities listed for them, and rated degree of confidence in their diagnosis • blinding to other clinical data not explicitly indicated • gold standard: diagnosis as determined by follow-up imaging or by patient's clinical course • spatial resolution of digital images: 1760 x 2140 pixel resolution of initial digital image, printed without loss or compression to film or displayed on a 1024 x 1536 pixel monitor	**Image Analysis** • no significant difference in accuracy for the detection of pneumothorax among three formats: small-format and large-format digital images printed on film, or digital images viewed on workstations • interpretation of digital images viewed on workstations was significantly less sensitive, but more specific, for the diagnosis of a pneumothorax than interpretations made from either format of digital image printed to film **Authors' Comment** • pneumothorax was chosen as a model because it appears as a subtle linear opacity at the margin of the lung, and it represents the type of subtle, high frequency abnormality that digital technologies must be able to detect in clinical settings ***Additional study findings reported in other tables***
Straub et al., 1991 *Department of Diagnostic Radiology, University of Pittsburgh*	**Purpose** to compare the diagnostic accuracy of interpreting abdominal mass CTs using digital images displayed in four modes: original CT film with 12 images/film; images displayed on 2 workstations in same pattern as on CT film; the same images, displayed sequentially on workstation; and images magnified x 2, displayed sequentially on the workstation **Cases/Controls** • 166 abdominal CT scans selected retrospectively: - 103 with abdominal masses (73 typical, 30 subtle) - 66 normal • total number of images of masses in any given abdominal organ ranged from 1 (prostate and musculoskeletal) to 54 (liver) **Methods** • images digitally-acquired using 512 x 512 x 12 bit format, and were available for review in each of four modes listed above (14" x 17" film + three styles of soft copy) • 8 radiologists interpreted all studies in each of 4 modes, with at least 5 weeks between viewing the same image in different modes • training handbook was given to each reader, to describe task and define abnormalities • order of individual radiologists' sessions and image presentation randomized • reading time not restricted. Use of window (contrast) and level (brightness) encouraged, but not quantified • readers identified abnormalities and rated both image quality and their level of comfort with the display mode • readers partially blinded; abnormalities to be rated in the study pre-defined during training • gold standard: verification by surgical report/ biopsy, CT or other imaging, or expert opinion based on clinical data • spatial resolution of monitor: 1536 x 2048 x 8 bits	**Image Analysis** • no statistically significant difference in accuracy for the detection of abdominal masses between digital images printed to film or viewed on a workstation in any of the configurations tested • readers reported the greatest comfort with the diagnostic task when viewing film, and the least comfort when viewing small sequential images on a workstation **Authors' Comments** • study demonstrates that soft copy images of CT can be used in this setting for the primary diagnosis of patients, since it can yield diagnostic results comparable to those obtained with film • assessment of comparable diagnostic performance is only a portion of the tasks that must be performed if PACS is to be successful in the clinical environment ***Additional study findings reported in other tables***

Study	Patients/Methods	Results/Comments
Razavi et al., 1992 *University of California School of Medicine, Los Angeles*	**Purpose** to compare the diagnostic accuracy of interpreting pediatric chest images using digital images printed as hard copy (film) vs. soft copy (on workstations) **Cases/Controls** • 239 neonatal lung images selected to represent a range of conditions which require high resolution imaging for diagnosis: - 162 images containing one or more abnormality of interest (26 pneumothorax,130 interstitial disease, 32 linear atelectasis, 88 air bronchograms) - 77 images with no tested abnormality • all selected images were of acceptable diagnostic quality **Methods** • computed radiography images captured in a 2048 x 2048 x 8 bit format • images were printed to three sizes of film. Films selected for study included 157 8" x 10", 44 10" x 12", and 38 14" x 17" images • readers were 3 senior and 2 junior pediatric radiologists trained in use of workstations, study forms, and use of defined diagnostic criteria for selected abnormalities • all readers reviewed all images, with 3-5 months between viewing the same image in different modes (hard vs. soft copy). Images were presented in random order • readers identified abnormalities, rated degree of confidence in their diagnosis, and recorded both viewing time and use of image-manipulation tools on workstation • readers partially blinded; diseases of interest pre-defined during training • gold standard: consensus opinion of 2 experienced pediatric radiologists based on all clinical and imaging data • spatial resolution of digital images: 2560 x 2048 pixels	**Image Analysis** • no significant difference in accuracy for the detection of pneumothorax and air bronchograms between hard copy (film) of digitized images and soft copy viewed on monitor • for all readers combined, "a slight performance edge " (not statistically significant) of workstations vs. film for the detection of interstitial disease and linear atelectasis on soft copy • for senior pediatric radiologists, workstations significantly more accurate than digital hard copy (size unspecified) for the diagnosis of interstitial disease and linear atelectasis **Authors' Comment** • the inherent smaller pixel size, and hence higher spatial resolution, of small plates used to produce hard copy images of pediatric films may account for differences between these findings and those of Cox et al. Results of this study may not be generalizable to adult chest imaging, where similarly high resolution hard copy might not be generated ***Additional study findings reported in other tables***

Table A2.4: Grading the quality of diagnostic accuracy studies: comparisons among conventional film, digital images displayed on workstations, and digital images laser printed to film

Notes:

- *Details of these studies are presented in Tables A2.1 through A2.3. All studies are observational, and most of them randomized the images shown to the readers.*
- *Operating characteristics are reported for each mode of image display analyzed in a study.*
- *Methods of analysis varied, and findings, including measures of variability when available, are reported in the form in which they were published.*
- *Studies in which a firm gold standard was present (physical evidence, final diagnosis incorporating best clinical judgement including follow-up exam) were rated as "+". For studies using a less definitive gold standard, and hence more susceptible to bias, the reference standard used is noted in the table.*
- *Most studies included in this table were small, and lacked the power to determine differences in accuracy among the imaging modalities, if they truly existed.*
- *No studies analyzed the observed effect size (difference in operating characteristics among different imaging technologies) relative to observed inter- or intra-observer variability.*
- *Abbreviations are listed at the end of this table, and terms are defined in the Glossary.*

** = statistically significant difference from conventional imaging* *# = statistically significant difference between hard copy (film) and soft copy (workstation) of digital image*

Study	Number	Measure of Diagnostic Accuracy			Evidence-Based Medicine Criteria				Quality Grade
		Conventional Film	Workstation Image	Laser-printed Film of DIgital Image	Comparison group	Gold standard	Blinding test reader to diagnosis	Blinding diagnosing MD to test results	
Slasky et al., 1990	187 abnormal lungs 113 controls (total)								
	62 interstitial disease:	Az=.89	Az=.87*	Az=.87					
	44 nodules:	Az=.86	Az=.85	Az=.85	+	+	+	+	B
	34 pneumothorax:	Az=.97	Az=.93*	Az=.93*					
	47 multiple diseases:	Az=.88	Az=.85*	Az=.87					
Cox et al., 1990	99 abnormal lungs 64 controls (total)								
	34 costophrenic angle blunting:	Az=.890 (SD=.031)	Az=.934 (SD=.020)	Az=.907 (SD=.022)					
	32 interstitial disease:	Az=.919 (SD=.019)	Az=.838 (SD=.034)*	Az=.884 (SD=.021)#					
	28 atelectasis:	Az=.871 (SD=.023)	Az=.888 (SD=.024)	Az=.880 (SD=.025)					
	26 pneumothorax:	Az=.982 (SD=.010)	Az=.898 (SD=041)*	Az=.981 (SD=.013)#	+	analog film	+	+	C
	23 parenchymal mass:	Az=.863 (SD=.046)	Az=.951 (SD=.020)	Az=.956 (SD=.014)*					
	22 consolidation:	Az=.806 (SD=.046)	Az=.789 (SD=.044)	Az=.805 (SD=.053)					
	17 obstructive disease:	Az=.857 (SD=.034)	Az=.797 (SD=.052)	Az=.892 (SD=.025)#					
	13 hilar/mediastinal mass:	Az=.826 (SD=.052)	Az=.842 (SD=.048)	Az=.913 (SD=.022)					
	7 apical scarring:	Az=.918 (SD=.033)	Az=.910 (SD=.045)	Az=.902 (SD=.038)					

Study	Number	Measure of Diagnostic Accuracy			Evidence-Based Medicine Criteria				Quality Grade
		Conventional Film	Workstation Image	Laser-printed Film of DIgital Image	Comparison group	Gold standard	Blinding test reader to diagnosis	Blinding diagnosing MD to test results	
Paakkala et al., 1991	146 chest images, 38% of them abnormal: 226 bone images, 58% of them abnormal:	note: small pleural or minor degenerative changes excluded from analyses Se=94.2% Sp=97.1% Accuracy=96% Se=94.3% Sp=91.5% Accuracy=94.6%	no statistical tests of significance reported Se=89.3% Sp=91.5% Accuracy=89.9% Se=89.5% Sp=91.5% Accuracy=89.9%		+	+	+	–	C
Elam et al., 1992	23 pneumothorax (range of severity) 22 age-matched controls (no disease or no pneumothorax)	Se=82% Sp=91% Az =.915 (95% CI=.819-1.011)	Se=65 % (NS) Sp=96% Az=.869 (95% CI=.755-.983)	small format film: Se=78% Sp=92% Az=.875 (95% CI=.759-.991) large format film: Se=77% Sp=90% Az=.874 (95% CI=.754-.994)	+	+	+	+	C
Franken et al., 1992	newborn chest: 4 edema, 4 BPD 8 HMD, 9 emphysema 10 pneumonia 14 normal newborn abdomen: 2 air in peritoneum 2 incarcerated hernia 3 obstructed intestines 16 necrotizing enterocolitis 19 no disease	overall accuracy, by reader: #1 Az=.8611 #2 Az=.8561 #3 Az=.8512 #4 Az=.8187	#1 Az=.8641 #2 Az=.8929 #3 Az=.8777 #4 Az=.8348		+	+	+	+	C
Ackerman et al., 1993	40 fractures 40 controls 40 pneumonias 40 controls	overall: Se=78% Sp=96% Accuracy=92% Az=.9475	Se=89%* Sp=92% Accuracy=84%* Az=.8765*		+	analog film	+	partial	C

Study	Number	Measure of Diagnostic Accuracy			Evidence-Based Medicine Criteria				Quality Grade
		Conventional Film	Workstation Image	Laser-printed Film of DIgital Image	Comparison group	Gold standard	Blinding test reader to diagnosis	Blinding diagnosing MD to test results	
Goldberg et al., 1993	adult: 196 abnormals 334 normals (total)	NA, since film was used as gold standard. However, authors noted that, when reviewing discrepant readings, expert panel identified 3 errors in film readings, 1 false negative and 2 false positives	overall adult: Se=96% (95% CI=93%-98%) Sp=99% (95% CI=98%-100%) accuracy=98% (95% CI=97%-99%) FP=1.0% FN=2.4%		+	analog film + consensus panel for discrepant cases	partial	-	D
	145 bone:		Se=97% FP=0.0% Sp=100% FN=2.5%						
	44 chest:		Se=95% FP=4.5% Sp=98% FN=1.8%						
	5 KUB:		Se=83% FP=0.0% Sp=100% FN=6.7%						
	2 sinus:		Se=100% FP=0.0% Sp=100% FN=0.0%						
	pediatric: 36 abnormals 119 normals (total)		overall: Se=94% FP=5.6% Sp=98% FN=1.7%						
	145 bone:		Se=100% FP=7.1% Sp=99% FN=0.0%						
	44 chest:		Se=93% FP=0.0% Sp=100% FN=3.7%						
	2 sinus:		Se=88% FP=12% Sp=90% FN=10%						
Thaete et al., 1994	133 single + 65 multiple chest abnormalities 112 normals (total)				+	+	+	+	**C** for rib fractures (under-powered) **B** for other disorders
	24 alveolar infiltrate:	average Az=.87 (range=.83-.92)	average Az=.84* (range=.77-.90)	average Az=.86 (range=.76-.89)					
	124 interstitial disease:	average Az=.77 (range=.72-.79)	average Az=.74* (range=.68-.78)	average Az=.75* (range=.70-.78)					
	84 nodules:	average Az=.84 (range=.82-.87)	average Az=.82 (range=.78-.86)	average Az=.84 (range=.82-.87)					
	27 pneumothorax:	average Az=.94 (range=.86-.99)	average Az=.86* (range=.78-.92)	average Az=.94* (range=.84-.98)					
	18 rib fractures:	average Az=.85 (range=.72-.95)	average Az=.85 (range=.76-.88)	average Az=.86 (range=.75-.95)					

Study	Number	Measure of Diagnostic Accuracy			Evidence-Based Medicine Criteria				Quality Grade
		Conventional Film	Workstation Image	Laser-printed Film of DIgital Image	Comparison group	Gold standard	Blinding test reader to diagnosis	Blinding diagnosing MD to test results	
Korsoff et al., 1995	25 abnormal lungs 16 controls (total) 15 subtle pneumothorax: 12 interstitial disease:	 average Az=.999 (range: .999 to 1.0) average Az=.878 (range: .845 to .947)	 1K2 average Az=.928* (range .836-1.00) 2K2 average Az=.983 (range .960-1.00) 1K2 average Az=.877 (range .832-972) 2K2 average Az=.831 (range .674-.895)		+	+	partial	+	C
Scott et al., 1995	60 selected abnormalities (musculoskeletal, abdominal, chest) 60 controls	all cases, all MDs: Se=50.2% Sp=79.2% Accuracy=64.5% all cases, ER MDs only: Se=38.1% Sp=72.0% Accuracy=54.8% all cases, radiologists only: Se=62.3% Sp=86.4% Accuracy=74.2%	Se=39.6%* Sp=75.6% Accuracy=57.3%* Se=29.1% Sp=69.9% Accuracy=49.2% Se=50.0%* Sp=81.4% Accuracy=65.4%*		+	analog film, with discrepant cases reviewed by expert panel	partial	+	C
DeCorato et al., 1995	812 images: 693 conventional film, 118 CT, 1 MRI 436 of these were normal (image type not specified)	NA, since film was used as gold standard. However, authors noted that, when reviewing discrepant readings, expert panel identified 2 errors in film readings	overall: 95% (95% CI=93.6%-96.6%) of workstation images not clinically significantly different from film image		+ (internal controls)	analog or digital film, with discrepant cases reviewed by expert panel	+	?	D
Kondoh et al., 1994	20 subtle interstitial lung lesions 20 controls (total) 8 collagen diseases 5 hypersensitivity pneumonitis or panbronchiolitis 1 metastatic calcification 1 sarcoidosis 1 alveolar proteinosis 1 eosinophilic granuloma 2 interstitial lesions of unknown origin	all diseases, all (11) readers: Az=.749 (SEM=.016) experienced radiologists (7): Az=.889 (SEM=.005) residents (4): Az=.812 (SEM=.029)		2/3-sized film: Az=.833 (SEM=.018) full-sized film: Az=.830 (SEM=.013)* 2/3-sized film: Az=.843 (SEM=.016)* full-sized film: Az=.847 (SEM=.014)* 2/3-sized film: Az=.816 (SEM=.021) full-sized film: Az=.800 (SEM=.034)	+	CT (see EG note)	-	?	D

Study	Number	Measure of Diagnostic Accuracy			Evidence-Based Medicine Criteria				Quality Grade
		Conventional Film	Workstation Image	Laser-printed Film of DIgital Image	Comparison group	Gold standard	Blinding test reader to diagnosis	Blinding diagnosing MD to test results	
Yoshino et al., 1992	25 cervical spine fractures 25 controls (5 normal, 20 with degenerative disc disease)	overall accuracy, by reader: #1 Az=.964 (SD=.030) #2 Az=.908 (SD=.043) #3 Az=.907 (SD=.053) #4 Az=.837 (SD=.069)		#1 Az=.883* (SD=.055) #2 Az=.901 (SD=.044) #3 Az=.840* (SD=.064) #4 Az=.847 (SD=.062)	+ (included only 5 true normals)	+	?	+	C
Straub et al., 1991	CT scans: 103 abdominal masses 66 controls (total) 6 pancreas 9 adrenal 5 spleen 54 liver 4 kidneys 4 retroperitoneal 4 lymph nodes 1 prostate 9 adnexa 2 uterus 4 bowel 1 musculoskeletal		overall accuracy (8 readers): images presented in mosaic pattern: Az=.90 (range=.88-.94) small images presented sequentially: Az=.92 (range=.91-.94) enlarged images presented sequentially: Az=.91 (range=.88-.95)	Az=.91 (range=.85-.96)	+	+	partial	+	C
Razavi et al., 1992	239 abnormal pediatric lung images 77 controls (total) 26 pneumothorax: 130 interstitial disease: 32 linear atelectasis: 88 air bronchogram:		Az=.8888 (SD=.075) Az=.8675 (SD=.028) Az=.8863 (SD=.058) Az=.7803 (SD=.072)	Az=.9079 (SD=.060) Az=.8459 (SD=.029) Az=.8349 (SD=.075) Az=.7770 (SD=.074)	comparison group contained "no tested abnormalities"	+	+	partial	C

Abbreviations:

Az, area under ROC curve
CI, 95% Confidence Interval
BPD, bronchopulmonary dysplasia
FP, false positive
FN, false negative
HMD, hyaline membrane disease
Range, lowest to highest Az values across readers
Se, sensitivity
SEM, standard error of the mean
Sp, specificity

Appendix 3

Evidence Tables for Clinical and Production Efficiency Studies

Table A3.1: *Summary of the literature*
Clinical and production efficiency in image management: comparison of film to PACS workstation environments

Notes:

- *Studies included in this table represent the strongest evidence presently available for the assessment of the efficiency of selected activities in a PACS workstation environment relative to a film-based environment.*
- *All are observational studies using mixed study designs including elements of case reports and case series.*
- *The study by Franken assessed the diagnostic accuracy of conventional film relative to that of digital imaging viewed on a workstation; Lou and Huang provided only preliminary impressions of the diagnostic accuracy of digital imaging viewed on a workstation relative to film imaging. Studies by Horii, Kundel, and Gay were designed to assess process effectiveness. They assumed, but did not assess, comparability of diagnostic accuracy of workstation imaging and film imaging. This assumption is not fully supported by available data.*
- *While the studies do suggest process and outcomes findings which are valuable for future assessments of the effectiveness (and hence the cost-effectiveness) of PACS, their methodological limitations should be considered when interpreting findings.*

Study	***Patients/Methods***	***Results/Comments***
Franken et al., 1992 *Department of Radiology, University of Iowa*	**Purpose** • to compare the diagnostic accuracy of interpreting clinical neonatal radiographs using conventional film images vs. digital workstation • to compare observer interpretation time for experienced readers using digital workstation vs. conventional film viewed on a workstation **Cases/Controls** • case sample of 100 chest or abdominal films from neonatal ICU, representing a wide range of subtle manifestations of common pediatric diseases - 58 chest images, 14 of which were normal - 42 abdominal images, 19 of which were normal **Methods** • film images were digitized for viewing on a PACS workstation • 4 radiologists read images, and ranked degree of confidence in their conclusions • 2 radiologists blinded to patient age, 2 aware of patient age • gold standard: confirmed diagnosis based on demonstrated typical clinical course of disease; subsequent imaging reflecting fulminating disease; or alternative diagnostic testing with a higher degree of accuracy • spatial resolution of digital images: 1024 x 1024 pixels. Small pediatric images could be displayed at full resolution, with imaging comparable to that of a 2000 x 2000 pixel monitor • image access time and image interpretation time study only included data for the 2 radiologists who were experienced with use of workstations • time elements measured by direct observation, using a stopwatch (time and motion study) • time to access PACS images: time to access images in local memory • time to access films on alternator was measured using comparison data gathered during a conference and a normal reading session. Films of study population were mounted in the same sequence as they were read, therefore time to locate film has no real meaning. • time to find cases in the workstation file menu, time to locate films in the file room, and time to mount films for viewing on the alternator were not measured • frequency of use and perceived usefulness of windowing, leveling, and magnification options was collected for the same 2 radiologists	**Image Acquisition TIme** • no significant difference in image acquisition time between PACS images and film mounted on alternators **Image Interpretation Time** • mean interpretation time for PACS digital images significantly longer than for interpretation of film images (36 and 47 seconds for each of interpreters using PACS vs. 22 and 20 seconds for each of interpreters using film) **Accuracy of Image Interpretation** • “no appreciable difference” in accuracy of interpretation between plain-film and digitized image viewed on workstation **Other Findings** • no difference in the general finding for specific diseases between two observers - windowing (contrast control) and leveling (brightness control) “almost always used” by both radiologists, and was considered useful in 80% of cases - magnification use varied greatly (67% vs. 25%) as did assessment usefulness of magnification (useful in half the cases vs. rarely useful) **Authors’ Comment** • a large part of the extra time in image interpretation of PACS images is related to the use of windowing, leveling, or magnification options

Study	Patients/Methods	Results/Comments
Horii et al., 1992 *Georgetown University Hospital, Washington, D.C.*	**Purpose** to compare case retrieval times for films stored in file room and images archived in a PACS **Cases** • retrospective case series of 70 randomly selected ultrasound and CT images: - 40 cases from current to 6 months old (30 US, 10 CT) - 20 cases from the past year (8 US, 12 CT) - 10 cases more than 1 year old (5 US, 5 CT) • the same cases were retrieved from the film file and from the PACS **Methods** • film storage process: - current studies maintained in Active File for 1 month, moved to Master File for 3 months, then to Archive for 1 year, finally to long-term off-site storage • PACS image storage process: - current studies automatically routed to workstations for review and also stored on magnetic disks in the database management system (DMS) for 5 days; then moved to optical disk storage in jukebox for 1 year; optical disks then shelved for long-term storage • image retrievals were performed at random times during the work day • time measurements recorded by the study personnel who submitted requests for images to hospital personnel, who were not blinded to study goals • time to retrieve film: measured from when the person got on line in the file room and stopped when the film was retrieved and verified to be correct • time to retrieve PACS image: measured from when the patient name or medical record number was entered on the workstation and stopped when the image appeared on the display was verified to be correct	**Image Retrieval Time** • overall, mean retrieval time for film statistically significantly longer than for PACS images: - PACS: 5 minutes, 17 seconds (SD not reported) - film: 6 hours, 48 minutes, 40 seconds (SD not reported) • no statistically significant difference between mean retrieval time for film vs. PACS images for cases ≤ 6 months old - PACS: 3 minutes, 26 seconds (SD ± 1 minute, 58 seconds) - film: 6 minutes, 7 seconds (SD ± 10 minutes, 38 seconds) • mean retrieval time for film statistically significantly longer than for PACS images for cases > 6 months old - PACS: 3 minutes, 1 second (SD ± 39 seconds) - film: 1 hour, 18 minutes, 14 seconds (SD ± 1 hour, 21 minutes, 18 seconds) **Image Retrievability** % of Cases Performed but Not Retrievable (lost) Study age / PACS / Film ≤ 6 months / 5 % / 10 % > 6 mo and ≤ 1 yr / 15 % / 30 % > 1 year / 30 % / 40 % **Authors' Comments** • delays in retrieving prior images can delay the performance or the interpretation of new imaging exams • requests for the retrieval of old exams decline exponentially with the age of the exam • there was no true data loss with PACS. Failure to find image was caused by errors when technologists captured images into the PACS system

Study	Patients/Methods	Results/Comments
Lou and Huang, 1992 *University of California, Los Angeles*	**Purpose** to assess efficiency of a PACS with display workstations relative to a film-based system for image delivery, system availability, and user acceptance in a neuroradiology setting **Cases (for study of image delivery and retrieval time)** • prospectively studied 1.5 years of imaging, 4984 CT and 4321 MRI examinations • 40% of exams accessed for viewing on PACS workstation **Methods** • PACS module, with 1280 X 1024 pixel display workstations, installed in neuroradiology reading room. Film alternators and light box remained available for film reading • radiologists free to view images on PACS workstation or as film • time measurement data were self-recorded or self-reported • film management process: - technologist at CT/MR created film from digital image - librarian delivered films to neuroradiology office - clerk retrieved patient's prior films - resident or librarian delivered films to reading area • PACS image management process: - computer received image from CT/MR scanner and formatted into image file - image transferred to data management computer - image transferred to display workstation • PACS performance measurement: - computer automatically logged time of each process and size of file - downtime of each component measured for 2-month period. Probability of being able to perform a particular task was calculated by multiplying together the % of time each piece of equipment needed to perform the task was functioning • film image performance measurement: - each film management step was "timed and estimated" by "several" experienced technologists, film clerks, and other personnel - retrieval of films from long-term storage and lag time between film pickup and delivery were not timed • PACS diagnostic accuracy and user acceptance pilot study: - evaluated by observation, survey of 10 users, and by preliminary impressions of image quality reported by 2 faculty and 1 fellow	**Image delivery times** • average image delivery time of the PACS system (88.5 minutes for CT, 44.2 minutes for MRI) is shorter than the average delivery time of the film-based system (93 minutes), even though archived films were not retrieved **Image retrieval times** • archived images can be retrieved from the PACS in an average of 2.9 minutes (CT) or 3.75 minutes (MRI) vs. an average of 15 minutes for retrieval of recent film **Image availability** • display workstation with image processing capacity available 99.97% of time. Any images stored in workstation can therefore be viewed 99.97% of time • PACS-archived image available for viewing 95% of time • percentage of times requested films are available for viewing not reported **User acceptance** • overall PACS system performance rated 3.4 (scale: 1 = poor to 4 = excellent) • image manipulation functions (contrast, brightness, magnification etc.) which responded instantaneously, were used for viewing 25% of images read on PACS workstation • 40% of patients exams performed during study period were accessed for viewing on the PACS workstation. No data reported as to whether or not same studies were also read as hard copy (film) • while diagnostic accuracy was not explicitly measured, preliminary impression of users was that PACS workstations and film were comparably accurate for performing diagnostic tasks **Authors' Comments** • in most clinical situations, the latest previous examination is used to make a diagnosis • based on finding about time of image delivery, display quality, and reliability of the PACS system tested, system upgrades and infrastructure changes were made, including the installation of 2K x 2K x 16 bit image workstations

Study	Patients/Methods	Results/Comments
Kundel et al., 1996 *University of Pennsylvania Health System*	**Purpose** to compare production and clinical efficiencies of a film-only system (either conventional or hard copy of computed radiology images) with those of a PACS **Cases** (for the prospective study of analysis of time to initiate action) • 2 data sets were used: all chest films of a random cohort of 5 MICU patients plus a random sample of all non-routine bedside chest examinations from medical intensive care or medical intermediate care unit. These films were usually ordered to assess position of tubes and catheters (PTC, N = 386) or pulmonary and pleural problems (PPP, N = 471) • study period: 5 randomly selected days in each week of the nine 4-week data collection periods **Methods** • data collection schedule: - analog film baseline data collection (*5 months to install and learn use of CR)* - 1st PACS data collection(film remained available for viewing) - CR film data collection (PACS workstation turned off, and all CR images printed to film) *(one month for staff training on workstations)* - 2nd PACS data collection (workstation used for primary viewing of 94% of images) • spatial resolution of monitors: 2048 x 2560 x 8 bits • time measurement data were collected using computer records, time and motion studies using video cameras and direct observation, activity sampling, and self-reporting • clinical action triggered by imaging findings abstracted from medical records. If imaging findings resulted in decisions to take no new action, no data were available • efficiency of image management system: measured as the time to image delivery. Included time to process image, to retrieve prior studies, and to mount images on a viewer or to transmit them to a remote workstation • time to encounter imaging information: measured as the time from exam completion to the time at which physician first obtained information about exam (by viewing image, by oral report from radiologist, or by written radiology report) • primary interpretation of images by radiologist based on reading film images	**Efficiency of Image Management** (Image delivery time + Image retrieval time) • median time necessary to make image available for viewing was significantly shorter for PACS workstation (10 minutes) than for film (1 hour) • 75% of images available on workstations within 20 minutes of completion; 75% of images available on film within 1.8 hours of completion **Time to Encounter Imaging Information** • overall, no statistically significant difference in the time to use imaging information between film and PACS workstation imaging • time to access imaging information significantly shorter for studies of position of tubes and catheters than for studies of pulmonary and pleural problems • for studies of position of tubes and catheters, no statistically significant difference in time to encounter imaging information between film and workstation imaging • for studies of pulmonary and pleural problems, time to encounter image information was significantly shorter for workstations than for film during the first PACS data collection period. However, this time to encounter PPP image information on a workstation was not significantly shorter than the comparable time interval measured at baseline **Time to Initiate Clinical Action** • elapsed time between completion of exam until primary clinical action taken significantly shorter when workstation used for primary image interpretation than when analog film used for primary interpretation (median 2.5 vs. 4.4 hours) **Use of Physician Resources** • radiologists used as the first source of information significantly less frequently when images were viewed on PACS workstation than when images were viewed as film (26%-32% of time for workstation vs. 90%-92% of time for film) **Authors' Comments** • overall, physician workflow pattern did not change to match the large decrease in time to image availability when PACS was installed • baseline time to take action after imaging to assess tube and catheter placement may have been the minimal clinical reaction time possible within the unit's work flow structure • when both film and workstation images available, physicians tended to use workstation images in acute situations (critical illness, admissions, line or tube placement) • time to take positive action based on imaging information was used as a surrogate for patient outcome. However, decreased time to take action does not necessarily result in improved patient outcome • accuracy of diagnosis made by MICU physicians using workstations as compared with final diagnosis made by radiologists from film not assessed • the simultaneous availability of the radiologist's official report along with the image itself should "increase confidence in clinical decision-making"

Study	Patients/Methods	Results/Comments
Gay et al., 1997 *University of Virginia Health Sciences Center, Charlottesville, VA*	**Purpose** to compare production and clinical efficiencies of a film-only system (hard copy of computed radiology CT images) with those of a PACS using 2K x 2K x 8-bit workstations (soft copy) **Cases** • for prospective study of workflow analysis: 10 body CT scans • for prospective study of events occurring during the work process: 54 body CT scans • cases were randomly selected from the workday, and were a mix of in-patient and out-patients referred for CT of the chest, abdomen, or pelvis • case mix not reported, but findings did indicate that interpretation involved measurement of lesions for 24 cases **Methods** • full-body CT exams were performed, and were either laser-printed to film or transferred to a PACS for workstation viewing • workflow analyses were performed by an independent observer who monitored each step of the work process using a digital stopwatch • a resource table, listing the steps, the resources used, and the mean time per work step, was created for both film and workstation image interpretation • bottlenecks (rate-limiting steps) in the work processes were identified • categories of events which occurred during the image interpretation sessions were identified • no statistical analyses were performed to compare average times per work step in a film as compared with a workstation environment. Work process was modeled and evaluated with a mean value analysis using Little's law	**Image Retrieval Time** • the average time to retrieve and display old exams for workstation viewing was 1.838 minutes (for images stored on the network), and 8.755 minutes (for archived images) compared with 58.690 minutes to retrieve old films from storage **Image Interpretation Time** • the average time to read images was 7.108 minutes per scan for film, as compared with 4.205 minutes per scan for workstation images • 87% of the workstation readings involved image manipulation **Workflow Analysis** • the retrieval of film folders by the file room personnel was observed to be the bottleneck in the work process of image management when using film • in a PACS environment, the technologist conducting the patient exam was found to be the bottleneck in the work process **Authors' Comments** • the finding that PACS reading required less time than did film reading may reflect the relative greater experience of these readers with PACS workstations • because comparison to previous exams is critical for many follow-up CT studies, when both PACS and film are available, a major use of the PACS station is the retrieval of old studies when old films are unavailable • image interpretation is the most critical part of the path for patient care and radiologist efficiency. Work process scheduling should be altered to optimize throughput for interpretation

Table A3.2: *Summary of the literature*
Image interpretation time: comparison of film (analog or digital) images to PACS workstation imaging

Notes:

- *Studies included in this table represent the strongest level of evidence presently available regarding assessment of image interpretation times. All are observational studies which used either a case series or a case-control design.*
- *With the exception of Kato (1995), all were primarily designed to assess diagnostic accuracy. Kato (1995) assumed comparable diagnostic accuracy of film and workstation imaging, an assumption not fully supported by the literature.*
- *Only one study (Franken, 1992) assessed the frequency of use of image processing tools (windowing, leveling, magnification, etc.).*
- *Studies varied in systems architecture and methodology, but did include clear descriptions of the elements of the clinical process which were timed.*

Study	***Patients/Methods***	***Results/Comments***
Straub et al., 1991 *Department of Diagnostic Radiology, University of Pittsburgh*	**Purpose** to compare the diagnostic accuracy of interpreting abdominal mass CTs using digital images displayed in four modes: original CT film with 12 images/film; images displayed on 2 workstations in same pattern as on CT film; the same images, displayed sequentially on workstation; and images magnified x 2, displayed sequentially on the workstation **Cases/Controls** • 166 abdominal CT scans selected retrospectively: - 103 with abdominal masses (73 typical, 30 subtle) - 66 normal • total number of images of masses in any given abdominal organ ranged from 1 (prostate and musculoskeletal) to 54 (liver) **Methods** • images digitally-acquired using 512 x 512 x 12 bit format, and were available for review in each of four modes listed above (14" x 17" film + three styles of soft copy) • 8 radiologists interpreted all studies in each of 4 modes, with at least 5 weeks between viewing the same image in different modes • training handbook was given to each reader, to describe task and define abnormalities • order of individual radiologists' sessions and image presentation randomized • reading time not restricted. Use of window (contrast) and level (brightness) encouraged, but not quantified • image interpretation time recorded by computer: elapsed time from the presentation of an image until the "done" instruction was entered into the computer scoring form used by the reader • readers identified abnormalities and rated both image quality and their level of comfort with the display mode • readers partially blinded; abnormalities to be rated in the study pre-defined during training • gold standard: verification by surgical report/ biopsy, CT or other imaging, or expert opinion based on clinical data • spatial resolution of monitor: 1536 x 2048 x 8 bits	**Image Interpretation Time** • no significant difference in the average interpretation time for film vs. workstation, in any configuration. Differences in reading time within the range of intrareader variability • no significant correlation between image interpretation time and performance of reader **Image Analysis** • no statistically significant difference in accuracy for the detection of abdominal masses between digital images printed to film or viewed on a workstation in any of the configurations tested **Authors' Comments** • workstation images of CT can be used in this setting for the primary diagnosis, since it can yield diagnostic results comparable to film • assessment of comparable diagnostic performance is only a portion of the tasks that must be performed if PACS is to be successful in the clinical environment ***Additional study findings reported in other tables***

Study	Patients/Methods	Results/Comments
Razavi et al., 1992 *University of California School of Medicine, Los Angeles*	**Purpose** to compare the diagnostic accuracy of interpreting pediatric chest images using digital images printed as hard copy (film) vs. soft copy (on workstations) **Cases/Controls** • 239 neonatal lung images selected to represent a range of conditions which require high resolution imaging for diagnosis: 162 cases (pneumothorax, interstitial disease, linear atelectasis, air bronchograms) and 77 normal images **Methods** • readers were 3 senior and 2 junior pediatric radiologists trained to use workstations, study forms, and defined diagnostic criteria • all readers reviewed all images, with 3-5 months between viewing the same image in different modes. Images were presented in random order • readers identified abnormalities, rated degree of confidence in their diagnosis, and recorded both viewing time and use of image-manipulation tools on workstation • image interpretation time recorded by computer: elapsed time from the presentation of an image until the "done" instruction was entered into the computer scoring form by the reader • blinding of readers to clinical information not specified • gold standard: consensus opinion of 2 experienced pediatric radiologists based on all clinical and imaging data • digital image resolution: computed radiography images captured in a 2048 x 2048 x bit format and displayed on a 2048 x 2560 pixel monitor	**Image Interpretation Time** • no significant difference in mean interpretation time for film vs. workstation • no significant differences in mean interpretation times for junior vs. senior radiologists **Image Analysis** • no significant difference in accuracy for the detection of pneumothorax and air bronchograms between hard copy (film) of digitized images and soft copy viewed on monitor **Authors' Comment** • the inherent smaller pixel size, and hence higher spatial resolution, of small plates used to produce hard copy images of pediatric films may account for findings. Results may not be generalizable to adult chest imaging, where similarly high resolution hard copy might not be generated ***Additional study findings reported in other tables***
Franken et al., 1992 *Department of Radiology, University of Iowa*	**Purpose** • to compare the diagnostic accuracy of interpreting clinical neonatal radiographs using conventional film images vs. digital workstation • to compare observer interpretation time for experienced readers using digital workstation vs. conventional film viewed on a workstation **Cases/Controls** • case sample of 100 chest or abdominal films from neonatal ICU, representing a wide range of subtle manifestations of common pediatric diseases: - 58 chest images, 14 of which were normal - 42 abdominal images, 19 of which were normal **Methods** • film images were digitized for viewing on a PACS workstation • 4 radiologists read images, and ranked degree of confidence in their conclusions • 2 radiologists blinded to patient age, 2 aware of patient age • gold standard: confirmed diagnosis based on demonstrated typical clinical course of disease; subsequent imaging reflecting fulminating disease; or alternative diagnostic testing with a higher degree of accuracy • spatial resolution of digital images: 1024 x 1024 pixels. Small pediatric images could be displayed at full resolution, with imaging comparable to that of a 2000 x 2000 pixel monitor • image access time and image interpretation time study only included data for the 2 radiologists who were experienced with use of workstations • time elements measured by direct observation, using a stopwatch (time and motion study) • time to access PACS images: time to access images in local memory • time to access films on alternator was measured using comparison data gathered during a conference and a normal reading session. Films of study population were mounted in the same sequence as they were read, therefore time to locate film has no real meaning. • frequency of use and perceived usefulness of windowing, leveling, and magnification options was collected for the same 2 radiologists *Note: time to find cases in the workstation file menu, time to locate films in the file room, and time to mount films for viewing on the alternator were not measured*	**Image Interpretation Time** •mean interpretation time for PACS digital images significantly longer than for interpretation of film images (36 and 47 seconds for each of interpreters using PACS vs. 22 and 20 seconds for each of interpreters using film) **Accuracy of Image Interpretation** •"no appreciable difference" in accuracy of interpretation between plain-film and digitized imaged viewed on workstation **Other Findings** •no difference in the general finding for specific diseases between two observers - windowing (contrast control) and leveling (brightness control) "almost always used" by both radiologists, and was considered useful in 80% of cases - magnification use varied greatly (67% vs. 25%) as did assessment usefulness of magnification (useful in half the cases vs. rarely useful) **Authors' Comment** •a large part of the extra time in image interpretation of PACS images is related to the use of windowing, leveling, or magnification options ***Additional study findings reported in other tables***

Study	Patients/Methods	Results/Comments
Thaete et al., 1994 *Department of Radiology, University of Pittsburgh*	**Purpose** to compare the diagnostic accuracy of interpreting adult chest images using three formats: conventional film images, digital images presented as hard copy (film), and digital images displayed on a workstation **Cases/Controls** • 310 chest radiographs selected from outpatient facility: - 198 with one or more abnormalities - 112 with normal chest images **Study Design/Methods** • conventional film images and computed radiography (CR) images were obtained from the same patient within 1 minute of each other • spatial resolution of digital images: 4096 x 5000 x 12 bit digital images displayed on 1536 x 2048 monitor or printed to film with high-resolution printer • 9 board-certified radiologists (7 with extensive workstation experience and 2 who were trained prior to study) interpreted images • readers were blinded to clinical information • unrestricted time allowed for interpretation of each image • radiologists were encouraged to use windowing and leveling options • image interpretation time recorded by computer: elapsed time from the presentation of an image (in any format) until the "done" instruction was entered in computer by the reader *Note: time to find cases in the workstation file menu, time to locate films in the file room, and time to mount films for viewing not measured*	**Image Interpretation Time** •no significant differences in mean image interpretation time for the three image modalities: - conventional film: 75 seconds (range 43-102 seconds) - laser-printed CR film: 71 seconds (range 41-96 seconds) - workstation: 74 seconds (range 42-116 seconds) •the differences in average interpretation time for the three image types were within the range of the intra-observer variability **Accuracy of Image Interpretation** •overall accuracy in the detection of interstitial disease significantly higher for conventional film than for hard-copy (film) of digital image •no significant difference between conventional film and digital hard-copy for detection of alveolar infiltrate, nodules, pneumothorax, or rib fractures •overall accuracy in the detection of alveolar infiltrate, interstitial disease, and pneumothorax significantly higher for conventional film than for digital image soft copy (on workstation). •no significant difference between conventional film and workstation image for detection of nodules or rib fractures **Authors' Comments** •user-friendly workstation technology supported efficient reading of images •this study emphasizes the need for large, carefully-designed studies in this field ***Additional study findings reported in other tables***
Kato et al., 1995 *Toshiba Hospital, Tokyo, Japan*	**Purpose** to compare observer interpretation times for PACS workstations (CRT) vs. digital image hard-copy (film) read on a light box **Cases** • all radiographic images obtained during a 1 week period were randomly assigned to one of two groups: - 237 displayed on PACS workstation - 219 printed film and viewed on light box **Study Design/Methods** • random allocation of images to one of the two groups (film or PACS) • spatial resolution of CR image: 2048 x 2560, reduced to 1024 x 1024 for display on workstation (with the option to zoom and magnify) • 4 radiologists (2 certified, 2 trainees) read both film and workstation images • all were experienced in use of workstation • time elements measured by direct observation, using a stopwatch (time and motion study) • time to access PACS images: time to access images in local memory • image interpretation time for film (current and previous films were placed near light box along with medical record): the time required to place films on the light box, read medical record information, interpret image, report results • image interpretation time for workstation (current study and one prior study were pre-loaded in the workstation. If other images needed, they were retrieved from the file server): time to display image, adjust window setting, read medical record information, interpret image, report results *Note: accuracy of image interpretation not assessed*	**Image Interpretation Time** •no statistically significant difference in mean interpretation time between reading digital hard copy (film) vs. digital soft copy (on workstation): - film: 299 seconds (SD ± 228 seconds) - workstation: 307 seconds (SD ± 209 seconds) **Authors' Comments** •five factors may affect radiology department throughput: transfer time for order information, examination time (both essentially same for conventional and PACS environment), image transfer time from imaging equipment to reader, image interpretation time, time to provide report and images to referring physician (all may be influenced by PACS installation) •adjusting windowing to read MRIs introduced some delays in readings, but these were not statistically significant

Glossary:

Abbreviations and Definitions

GLOSSARY

**note that words in italics have been defined elsewhere in the glossary*

A_z: the area under the *ROC curve.* Used to compare the *accuracy* of diagnostic tests: the more accurate the test, the larger the area under the curve. This index varies from 0.5 (no apparent accuracy) to 1.0 (perfect accuracy).

Accuracy: the proportion of all test results (both positive and negative) that are correct; results close to the true measure of the biologic phenomenon. Accuracy depends on the *validity* and *precision* of the study.

Activity based costing (ABC): a cost accounting system developed by Cooper and Kaplan. ABC defines costs in terms of an organization's processes or activities and determines costs associated with significant activities or events. One effect of viewing costs within this framework is that much of what were thought to be *fixed costs* within a traditional accounting system are revealed to be *variable costs* associated with the volume of production of particular products and services.

Amortization: the process of paying off a debt liability and accrued interest through a series of equal, periodic payments.

Analog signal: a form of information transmission in which the signal varies in a continuous manner rather than in discrete steps.

Analytic perspective: the viewpoint chosen for an economic analysis (for example, that of the patient, the service, the institution, or society).

Architecture: the selection, design, and interconnection of the physical components of a computer system.

Bias: a type of systematic error. Bias can originate from many different sources, such as allocation of patients, analysis, interpretation, publication, and review of data. In the worst circumstances it may lead to the wrong conclusions being drawn.

Bibliographic database: an indexed computer or printed source of citations of journal articles and other reports in the literature. Bibliographic citations typically include author, title, source, abstract, and/or related information (including full text in some cases). Examples are MEDLINE and EMBASE.

Bit (Binary digit): the smallest piece of digital information that a computing device handles. It represents off or on (0 to 1). All data in computing devices are processed as bits or strings of bits.

Blinding: the concealment of group assignment (to either the treatment or control group) from the knowledge of patients and/or investigators in a clinical trial. *Blinding* eliminates the possibility that knowledge of assignment may affect patient response to treatment or investigator behaviors that may affect outcomes. *Blinding* is not always practical (e.g., when comparing surgery to drug treatment), but it should be used whenever it is possible and compatible with optimal patient care. A **single-blind** trial is one in which knowledge of group assignment is withheld only from patients; a **double-blind** trial is one in which the knowledge is withheld from patients and investigators.

Bottleneck: that resource that limits the upper bound on the throughput rate when performing a workflow analysis.

Case study (case report, anecdote): a type of nonexperimental (observational) study design in which the investigator reports an intervention and outcome in a single patient.

Case series: a type of nonexperimental study design in which an investigator reports a group or series of cases with the characteristic of interest. Although among the most common, case series are the weakest studies designed to establish causation.

Case-control study: a type of retrospective, nonexperimental study design in which individuals with a particular condition or disease (*cases*) are selected for comparison with a series of individuals in whom the condition or disease is absent (the *controls). Cases* and *controls* are then compared with respect to existing or past attributes or exposures.

Case: a person in the study group who has the disease or characteristic of interest.

Clinical significance: the effect that a *technology* or intervention has which is meaningful to patients and/or health care providers; however, it may or may not have *statistical significance*.

Cohort study: a type of nonexperimental study design in which outcomes are compared in a group of patients that received an intervention with a similar group (a cohort) of patients that did not.

Confidence interval (CI): depicts the range of uncertainty about the estimate of a parameter which has been calculated using the observations from a study. The true value of a parameter is thought to lie, with the specified level of confidence, within the confidence interval. The CI is related to the sample size used in a study. A small sample size provides less information than a large one, and the CI is correspondingly wider when a small sample is used.

Continuous variable: quantitative data that may take on fractional values (e.g., height, weight, serum cholesterol).

Control group: referent group; a group of study subjects to which the effects of an intervention given to the treatment group is compared and who, with the exception of the intervention, resemble the treatment group as closely as possible.

Cost-benefit analysis: an economic analysis which expresses the outcome of interest (or the benefit) in terms of currency (e.g., loss in net earnings due to death or disability).

Cost-effectiveness analysis: an economic analysis which compares the outcome of decision options in terms of their monetary cost per unit of health outcome achieved. Health outcomes are measured in terms of health status.

Cost-utility analysis: an economic analysis which incorporates relative social value or preferences into the health outcome considered. Outcomes are often expressed as a monetary cost per quality-adjusted life year.

CR (Computed Radiography): a storage phosphor plate contained in a cassette and used instead of a conventional film-screen cassette. A laser beam scans the exposed plate to produce the digital data that is then converted into an *image*.

Cross-sectional study: a type of nonexperimental study design in which a group is chosen (sometimes as a random sample) from a certain larger population, and the exposure of people in the group to an intervention and outcomes of interest are determined.

Database (register): any of a variety of repositories, often computerized for observations and related information about a group of patients, a disease, an intervention, or other events or characteristics.

Decision analysis: the modeling of the sequences of multiple possible strategies to determine which is optimal. It is based upon available estimates (drawn from the literature and/or from expert opinion) of the probabilities that certain events and outcomes will occur and the values of the outcomes that would result from each strategy.

Diagnosis: the process of determining one's health status and the factors responsible for producing it.

Diagnostic accuracy: a characteristic of *diagnostic test efficacy* describing the proportion of all test results that are correct.

Diagnostic test efficacy: the impact and usefulness of a diagnostic test expressed in terms of its technical properties, *diagnostic accuracy*, or its impact on *diagnosis*, therapy, patient *outcome*, or society.

DICOM (Digital Imaging and Communication In Medicine): a standard for interconnection of medical digital imaging devices, developed by a committee sponsored by the American College of Radiology and the National Electric Manufacturers Association.

Digital signal: a form of information transmission in which the signal varies in discrete steps, such as those represented by bits, rather than in a continuous manner.

Digitize: the process by which analog (continuous value) information is converted into digital (discrete value) information. This process is a necessary function for computer imaging applications because visual information is inherently in analog format and most computers use only digital information.

Direct file capture: the process by which *image* data is obtained directly from an *image* file. The *image* produced from the file, regardless of the *technology* that produced it (CT, MRI, CR, US) is identical to the original.

Direct image capture: the capture or acquisition of digital data that has already been recorded in digital form by an imaging modality.

Effectiveness: the extent to which an intervention produces favorable outcomes under usual or everyday conditions.

Efficacy: the extent to which an intervention produces favorable outcomes under ideal conditions, for example, within the protocol of a carefully managed randomized controlled trial, or at a "center of excellence."

Evidence table: a summary display of selected characteristics of studies of a particular issue of interest.

Evidence-based approach: the systematic location and critical appraisal of published research and other available literature.

Evidence-based clinical practice (EBCP): an emerging clinical discipline in which the best available evidence for research about *diagnosis*, prognosis, therapy, and other clinical and health issues is applied to decisions in health care.

Experimental study: a type of epidemiological study design in which the exposure or intervention of interest is assigned to study subjects by the investigator often in a randomized manner (e.g., randomized clinical trials) to reduce confounding; in evidence-based terms, this type of study provides stronger evidence supporting a casual link between the intervention and outcome(s) of interest.

Fixed costs: costs that do not vary with level of output or activity. Typically, these are considered overhead costs and are not included in a *cost-effectiveness analysis.*

Generalizability (external validity): the degree to which the results of a study hold true for situations other than those pertaining to the study, in particular, for routine clinical practice.

Gold standard: a method, procedure, or measurement which is widely conceived to be the best available, against which new interventions should be compared. It is especially important in the context of diagnostic testing.

Gray scale: the number of different shades of levels of gray that can be stored and displayed by a computer system. The number of gray levels is directly related to the number of bits used in each *pixel*: 6 bits = 64 gray levels, 7 bits = 128 gray levels, 8 bits = 256 gray levels, 10 bits = 1024 gray levels, and 12 bits = 4096 gray levels.

Gray-scale monitor: a black to white display with varying shades of gray, ranging from several shades to thousands, thus being suitable for use in imaging. This type of monitor also may be referred to as a monochrome display.

HIS (Hospital Information System): an integrated computer-based system to store and retrieve patient information including laboratory and radiology reports.

Hierarchy of evidence: study designs are often ranked according to their *validity*, or degree to which they are not susceptible to *bias*. The hierarchy indicates which studies should be given the most weight in a synthesis. Usually, well designed randomized clinical trials are seen as being at the top of the hierarchy, whereas observational studies or expert opinion are seen as lower down.

Image compression: method to reduce the amount of data needed to reproduce an *image*.

Image: a computer's representation of a physical object. For example, when the visual item is a photograph or radiograph displayed on the computer monitor or stored in a computer file.

K (Kilo): stands for the number one thousand. It is used primarily when referring to computer storage and memory capacities: for example, 1 Kybte = 1024 bytes.

Laser film scanner: a device that uses a laser beam to convert an *image* on x-ray into digital *image* data.

Little's law: an equation used in modeling and evaluating workflow. The equation states that the average number of jobs in a system is equal to the mean arrival rate of jobs to the system times the mean time for a job to flow through the system.

Mean: measure of central tendency describing the average value of a group.

Medical informatics: a field of study combining computer science, information science, and medicine which is concerned with a broad range of issues in the management, use and nature of biomedical information, and medical computing.

MEDLARS: Medical Literature Analysis and Retrieval System comprising about 40 computer databases managed by the National Library of Medicine.

MEDLINE: one of the most popular *MEDLARS* databases comprising bibliographic citations published since 1966 from about 3,700 health and biomedical journals.

MeSH: Medical Subject Headings. This is a controlled vocabulary of approximately 15,000 medical terms used to identify the subject content of the medical literature in *MEDLARS* databases.

Meta-analyses: a statistical analysis of the results of a collection of studies for the purpose of synthesizing their findings. These are particularly useful in summarizing prior research when individual studies are too small to yield valid conclusions.

Modem: a device that converts *digital signal*s from a computer to pulse tone signals for transmission over telephone lines.

Nonrandomized control (concurrent nonrandomized control): a control group that is observed by a research investigator at the same time as the intervention group, but that was not established using random assignment of patients to the control and intervention groups. Differences in the composition of the two groups may result.

Null hypothesis: a statement proposing that the intervention and the outcome being measured are not associated. This is the basis for the test of statistical significance. Statistical tests attempt to reject the null hypothesis of no association in favor of an alternative hypothesis that there is an association.

PACS: picture archiving and communication system.

Patient selection bias: error due to systematic differences between those who are included in the study and those who are not. This *bias* may affect the *external validity* of a study.

Pixel (picture element): the smallest piece of information that can be displayed on a computer screen. It is the fundamental picture element of a digital image, with the total *image* being composed of a large array of pixels.

Phase I, II, III, and IV studies: phases of clinical trials of new drugs in the drug development and approval process of FDA. Phase I trials are the first experiments in humans and are primarily concerned with establishing drug safety, metabolism, and dosage range. They are usually conducted on a small number of normal volunteers (depending on the nature of the drug and it's anticipated toxicity). Once safety is established, Phase II trials are then conducted on a larger number of volunteer patients. These are controlled trials conducted to determine efficacy and adverse reactions. Once a drug is shown to be reasonably effective and reasonably well-tolerated, full scale Phase III trials are conducted. These are large, rigorously designed clinical trials to

verify the efficacy and monitor adverse reactions that were observed in the earlier, more weakly-designed Phase II trial. Phase IV trials are postmarketing studies to monitor long-term effects.

This classification system has been used for general guidance for clinical trials not related to drugs, such as for assessing radiation treatment, and, to a lesser extent, for the assessment of new surgical techniques. It has recently been suggested that a similar classification system be used for the economic evaluations of promising health care interventions.

Power: the probability of rejecting a *null hypothesis* when the null hypothesis is indeed false; the relative frequency with which a true difference of specified size between the comparison groups would be detected by the intervention or test of interest.

Precision: the reproducibility of the study result, given similar circumstances. This can be affected by patient and laboratory conditions, inter-observer variation, and intra-observer variation.

Primary study: an investigation that collects original (primary) data from patients, e.g., *randomized clinical trials, observational studies, case series,* etc.

Resolution: the ability of an imaging system to differentiate between objects.

RIS: Radiology Information System.

Roam and zoom: the ability to select and magnify a region in the display.

ROC curve: receiver operating characteristic curve. A graphic means for assessing the ability of a diagnostic test to discriminate between disease and no disease. The term "receiver operating characteristic" comes from psychometry where the characteristic operating response of a receiver-individual to faint stimuli is recorded.

Sample size: the total number of subjects in a study, including both treatment and control groups.

Sensitivity: the proportion of people who truly have the disease who test positive for the disease.

Sensitivity analysis: a process through which the robustness of an economic model is assessed by examining the changes in results of the analysis when key variables are varied over a specified range.

Specificity: the proportion of people who truly are without the disease who test negative for the disease.

Statistical power: see *Power.*

Statistical significance: a conclusion determined by a statistical test that demonstrates whether a *technology* or intervention has a true effect on outcome over and above that which would have occurred by chance alone. Statistical significance does not prove causality nor does it provide information about the magnitude of the effect, nor is it sufficient to demonstrate the *clinical significance* of the *technology* or intervention on patient outcome.

Surrogate endpoint (or intermediate outcome): an outcome measure that is used in place of a primary endpoint (outcome). Examples are decrease in blood pressure as a predictor of decrease in strokes and heart attacks in hypertensive patients, and increase in T-cell (a type of white blood cell) counts as an indicator of improved survival of AIDS patients. The use of a *surrogate endpoint* assumes that it is a reliable predictor of the primary endpoint(s) of interest.

Technology assessment: any process of examining and reporting properties of a medical *technology* used in health care, such as safety, efficacy, feasibility, and indications for use, cost, and cost-effectiveness, as well as social, economic, and ethical consequences, whether intended or unintended. Its purpose is to support technology-related policy making in health care.

Technology: the drugs, devices, and medical and surgical procedures used in health care, and the organizational and supportive systems within which such care is delivered (Office of Technology Assessment, 1978).

Telemedicine: the use of electronic information and communications technologies to provide and support health care when distance separates the participants.

Total cost of ownership: a costing approach developed to capture the complete cost of owning and operating an information technology system. This approach uses classical accounting methods, and includes not only the vendor costs directly associated to the software and hardware, but the cost of maintaining and supporting the system, as well as the "productivity effects" of the system (Deloitte & Touche, 1997).

Validity (of a measurement): the degree to which a measurement truly measures what it purports to measure.

Validity (of a study): the degree to which the inference drawn from a study is justified. **Internal validity** is the degree to which the effect observed in a study can be attributed to the hypothesized effect under investigation. Internal validity is usually highest in large randomized controlled trials, and decreases with decreasing scientific rigor. **External validity** *(generalizability)* is the degree to which the results of a study hold true for situations other than those pertaining to the study, in particular, in routine clinical practice.

Variable costs: costs that change with the level of activity or output. These include the value of all those goods, services, and inputs in a product line or activity, and are traditionally included in a *cost-effectiveness analysis*.

Workstation: a functional grouping of computer hardware and software for individual uses such as word processing or image viewing.