

Peter Braun, MD
Nancy McCall, Sc.D.
RTI International

•
MedPAC

601 New Jersey Avenue, NW
Suite 9000
Washington, DC 20001
(202) 220-3700
Fax: (202) 220-3759
www.medpac.gov

•
The views expressed in this memo
are those of the authors.
No endorsement by MedPAC
is intended or should be inferred.

Improving the Accuracy of Time in the Medicare Physician Fee Schedule: Feasibility of Using Extant Data and of Collecting Primary Data

*A report by staff from RTI International for the Medicare
Payment Advisory Commission*

December 2011

December 2011

Improving the Accuracy of Time in the Medicare Physician Fee Schedule: Feasibility of Using Extant Data and of Collecting Primary Data

White Paper

Prepared for

**Kevin Hayes
MedPAC**

Prepared by

**Peter Braun, MD, Consultant
Nancy McCall, Sc.D.**
RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709

RTI Project Number 0210516.004.000

Contents

Section		Page
1.#	BACKGROUND AND USE OF EXTANT DATA	1-1#
1.1 #	Introduction and Study Objectives	1-1#
1.2 #	The Challenges and Merits of Replacing Subjective with Objective Time Data in the Medicare RBRVS Physician Fee Schedule (PFS)	1-2#
1.2.1#	RUC Recommendations on use of Extant Data	1-2#
1.3#	The Evolution of Methodology in Development of the RBRVS.....	1-3#
1.3.1#	Precedents in Medicare Physician Payment Policy and Present Challenges	1-5#
1.4#	Assessment of the Feasibility of using Extant Secondary Data Sources to Develop Objective Time Estimates	1-5#
2.#	THE FEASIBILITY OF COLLECTING PRIMARY DATA	2-1#
2.1#	Initial Telephone Interviews with Large Health Systems	2-1#
2.2#	Site Visits: Structure and Planning.....	2-1#
2.3#	Summary of Findings from Site Visits at System A and System B	2-2#
3.#	DISCUSSION AND RECOMMENDATIONS FOR SPONSORED STUDIES	3-1#
3.1#	Office-based E&M Services and Hospital E&M Services	3-1#
3.2#	Common E&M Services that may be Performed Differently by Different Specialties, including Non-Physician Providers	3-2#
3.3#	Office-based Surgical Procedures, Professional Component of Diagnostic Services, including Pathology Services, Imaging services	3-3#
3.4#	Surgical and Diagnostic Services Performed in Ambulatory Care Facilities ..	3-3#
3.5#	Major Surgical Procedures	3-4#
References		R-1#

Tables

Table 1. Summary of Findings from Site Visits Systems A and B..... 2-4

1. BACKGROUND AND USE OF EXTANT DATA

1.1 Introduction and Study Objectives

Research conducted by RTI for CMS and the Assistant Secretary for Planning and Evaluation (ASPE) has shown that some of Medicare's RBRVS fee schedule's estimates of physician time based upon surveys of physicians by specialty societies are considerably higher than estimates obtained from other data sources, (e.g., operating room logs) (McCall et al., 2006). That is, the subjective times on which both the work and practice expense RVUs for many **invasive services** are based may diverge significantly from objective measures of the time physicians take when performing these services.

There has long been evidence that there are significant differences in the times used by physicians in different specialties to perform **evaluation and management (E&M) services** coded the same, that work for certain services coded the same differs significantly across specialties and that paying the same for work that differs introduces serious distortions in payment (Braun et al., 1988). An historical perspective on this issue is provided in **Appendix A** to highlight the importance of obtaining time for E&M services across a broad array of specialties.

There are three objectives to this study:

- Identify and evaluate objective data currently available through extant databases on the time that physicians and non-physician clinical personnel spend in furnishing services billable under Medicare's physician fee schedule.
- Assess the feasibility of primary data collection that would provide objective time values for performance of surgical, diagnostic, including pathology and laboratory, and imaging services, across inpatient and ambulatory settings from a cohort of physician practices.
- Assess the feasibility of primary data collection that would provide objective time values for E&M services, which are characterized primarily by type of service (site, referral status, etc...) and level of service across a broad spectrum of different specialties from a cohort of physician practices.

In this White Paper, we summarize findings related to these three study objectives. We begin by discussing the properties of extant time data that the American Medical Association's RBRVS Update Committee (RUC) has advocated as being necessary before it can be incorporated into the updating of Medicare's physician fee schedule. Second, we provide an evaluation of currently existing databases' characteristics relative to the RUC recommendations for objective data. Third, we present a summary of our findings on the feasibility of primary data collection as assessed through site visits with two major healthcare organizations. Finally, we recommend a set of potential studies that would yield measures of objective time of physicians' evaluation and management, surgical, diagnostic, imaging and laboratory services.

1.2 The Challenges and Merits of Replacing Subjective with Objective Time Data in the Medicare RBRVS Physician Fee Schedule (PFS)

In the Third 5 Year Review, a number of specialty societies submitted surgical time estimates to the RUC from the National Surgical Quality Improvement Program (NSQIP) and the Society of Thoracic Surgeons' national data base (STS). The RUC used the data in a variety of ways – to validate survey time estimates or to raise survey time estimates (71FR No. 125). However, the RUC did express reservations about mixing data from different data sources and deviating from the accepted methodology of the RUC. In each instance that the RUC used the extant data, CMS rejected its use both on methodological grounds and concern about the characteristics of the databases. CMS expressed concern that the databases may not be “truly representative” and valid and that by applying the data randomly relativity between services could be distorted. In the Final Rule, CMS provided quotes from the defense of the NSQIP by the American College of Surgeons (ACS). In their comments, the ACS argued “the NSQIP data on intra-service time is verified through operating room logs and is the ‘absolute gold standard’ for estimating surgeon intra-service times; the best data should always be used rather than relying on ‘the lowest common denominator’; and there had never been a prior requirement that a single methodology be used to evaluate physician services across all services.” In response, CMS requested that the RUC and the specialty societies join CMS in further dialog concerning the role of extant databases.

1.2.1 RUC Recommendations on use of Extant Data

In February 2008, at the request of CMS, the RUC proposed the following exhaustive list of inclusionary and exclusionary criteria for determining the appropriateness of incorporating extant data in its updating of the RBRVS fee schedule¹.

1. *“Databases must have data integrity/reliability*
 - o *Must collect data prospectively,*
 - o *Should have the ability to identify and assess outliers – multiple procedures resulting in greater LOS; diseases with high mortality rate (LOS=0) or extended recovery (LOS>90); age variance (bi-modal)*
 - o *Should have the ability to have transparency of data to compare to other databases including the RUC database*
 - o *Should have the ability to audit the database*
 - o *Should have the ability to track the data/changes over time*
 - o *Should have the ability to collect data on all cases done by participants or for large volume procedures or E/M encounters, should have sampling criteria that are statistically valid to eliminate sampling bias*
 - o *Should have current data, preferably from the last three to five years, although older sets can be used for comparison purposes*

¹ The criteria are excerpted from a staff note on the RUC review of their Extant Data Policy.

2. *Databases should collect time data for the procedures, at a minimum the skin-to-skin or intra-service time and length of stay. Additional time elements may include ICU LOS, and other specialty specific time factors (e.g. phone calls, ventilator hours)*
3. *Must have the ability to unequivocally map the procedure to a CPT code and isolate the procedure from associated physician work that is otherwise billable in the same setting*
4. *Databases must list their limitations – include what is provided and not provided with respect to the RUC database*
5. *Databases must be representative*
 - o *The data should be geographically representative e.g., regionally and nationally for the specialty,*
 - o *The data should have various levels of patient severity*
 - o *The data should have adequate practice site representation and sample size – practice sites and rural and urban representation*
 - o *The data should be from various practice types – representative of the academic, non-academic and other types of practices for the specialty*
 - o *The data should be collected from the majority specialties (including subspecialties) that perform the procedure or encounter*
 - o *The data should be collected from either hospital/institution or individual physician.”*

Upon meeting all of the above listed criteria, the RUC approved the use of the extant data as supplemental data to the RUC survey process or as primary data in certain instances during the Five Year Review process.

We believe the collective set of criteria developed by the RUC creates an excessively high standard that no secondary source of objective data could meet. And, we do not believe that the time data collected by the specialty societies through their survey process meets the exacting requirements. As one example, an October 2008 Summary of Recommendations for the development of work RVUs for CPT code 32552, removal of indwelling tunneled pleural catheter with cuff, shows a survey response rate of only 8.4% with only 1% of responding physicians practicing in a rural area.

MedPAC has expressed concerns to CMS about the lack of information on some of the survey characteristics of the AMA's Physician Practice Information Survey (PPIS) that is the current source of practice expense data. In that letter, there is a request for more information regarding (1) response rate by specialty, (2) nonresponse adjustment method, (3) characteristics of survey responders to the entire specialty and non-responders, and (4) accuracy. These are not unreasonable criteria to apply to any dataset.

1.3 The Evolution of Methodology in Development of the RBRVS

The Harvard Investigators faced a difficult methodological challenge and logistical task in not only developing acceptable methods and establishing resource-based RVUs for more than 4,500 physician CPT codes between 1986 and 1992, when RBRVS became the basis for Medicare physician payment reform. Given the logistical challenge, it was not possible to obtain values for time other than through surveys of physicians. Statistical evaluation of

the survey data indicated that they had acceptable measures of internal validity. Testing of surveyed times for invasive procedures against a limited set of data for objective time (operating room log data) available at the time indicated that they had acceptable accuracy (external validity).

The investigators went forward with survey methods as the best that could be done at the time and under the circumstances. However, to meet the Health Care Financing Administration's (HCFA) timeline, *two different survey methods* were used to determine the original work relative values.

- The Harvard investigators used large national survey methods (n=160 physicians/specialty) for measuring work and time of approximately 60 services per specialty during 1986-1992 using magnitude estimation (Hsiao et al., 1988).
- The large number of remaining services, roughly 4,500 CPT codes, were measured by the small group methods (n=15 physicians per specialty) using the values of selected services determined in the national surveys as benchmarks (Leape et al., 1992).

We also note improvement in methods, aimed at having a more accurate correspondence of payment to resource inputs, has been given priority over mere "consistency" in the evolution of Medicare physician payment since 1986. Methods for work, practice expense, and professional liability insurance components of the RBRVS were, because of methods, needs for data, and effects on providers, all ***phased in over extended periods***.

- Resource-based work values were phased in over a five-year period (1992-1997) via a progressive mix of resource-based values and historic charge-based values.
- Introduction of resource-based practice expense values were delayed for seven years from the advent of physician-payment reform (1999) and phased in over a three year period (1999-2002).
- Introduction of resource-based professional liability insurance relative values were delayed for eight years (2000) and phased in over a two-year period (2000-2002).

Lastly, in CMS' response to its proposed rule for developing a formal process to validate relative value units under the PFS (FR Vol75 No228, Page73218), CMS writes "the AMA RUC does not rely on a single consistent methodology to value codes. Based on our historical and current review of the AMA RUC recommendation summaries...., we have noticed that the AMA RUC appears to use a variety of methodologies in its valuation process. For some codes, we have noticed that the AMA RUC uses magnitude estimation in conjunction with survey data from surveys conducted by the specialty societies to support the values. For other codes, the AMA RUC uses magnitude estimation to override the results of the survey data....The AMA RUC may also elect to use a crosswalk approach in valuing a code by applying a work value from a currently valued code under review based on the clinical similarity of the procedures....In some instances, we note that the AMA RUC has asserted that it uses the building block methodology to value the code."

1.3.1 Precedents in Medicare Physician Payment Policy and Present Challenges

The use of Harvard resource-based values for physician work in 1992 and introduction years later of resource-based values for practice expense and professional liability insurance are phases of an ongoing process to provide relative values that better correspond to resource inputs than the payment system in place before Medicare physician payment reform in 1992. The result is not perfect, but it is presumably better than what had come before. Policy makers have seen that in this arena the “perfect is the enemy of the good.” Requiring universal objective measures of time before changes are made in the PFS, as suggested by the AMA RUC, envisages a state of affairs that may take more resources and time than are likely to be available in the future and is inconsistent with the recent history of physician payment policy of the last two decades or more.

More recent comparison of both Harvard and RUC survey times with objective data from surgical operating room logs (DJ Sullivan) have shown systematic major upward deviation of survey data from objective times. The deviations in time were greater for the RUC surveys than the original Harvard times (McCall et al., 2006). In short, we need better data on surgical and other service times for an RBRVS that is valid and that provides a basis for more equitable payments among physicians.

We believe it is fully consistent with the history of Medicare physician payment policy to obtain objective measures of physician service time and to incorporate them, in a phased process, in the PFS. Thus, we believe it would be reasonable to consider the following principles for updating the Medicare RBRVS PFS:

- Methods for measuring objective time could be applied first to physician services that comprise a major portion of physicians’ services (and hence, physician payment) in each specialty and in overall Medicare payments. We note that 128 CPT codes accounted for two-thirds of Medicare Part B expenditures in 2008 (RTI analysis of 2008 high volume services spreadsheet provided by MedPAC staff).
- Over time, these methods would be applied to a larger set of services and ultimately to all or nearly all services.
- Data on objective time would be phased in (as with physician work, practice expense, and professional liability insurance costs) over reasonable periods of time.
- Resources adequate to this significant task would be made available by CMS or other governmental agencies and provided to independent investigators by competitive solicitation.

1.4 Assessment of the Feasibility of using Extant Secondary Data Sources to Develop Objective Time Estimates

We have undertaken a high-level evaluation of currently existing or extant databases and compared the attributes of the databases to the RUC criteria for extant data requested by MedPAC staff. We evaluated characteristics of the following six databases:

- Society of Thoracic Surgeons data base (STS)
- National Surgical Quality Improvement Program (NSQIP)
- National Ambulatory Medical Care Survey (NAMCS): Office
- National Ambulatory Medical Care Survey (NAMCS): Ambulatory Surgical Center (ACS)
- National Ambulatory Medical Care Survey (NAMCS): Emergency Room (ER)
- National Ambulatory Medical Care Survey (NAMCS): Hospital Outpatient Department (OPD)

While most of the evaluated databases possess many of the RUC's required properties, none possess all of them. The STS and NSQIP have the greatest level of specificity with respect to the physician service time. The NSQIP is the only database that explicitly links service time to individual CPT codes; a major limitation of all other databases. None of the individual databases span all types of services or places of service. While all databases provide transparency with respect to data collection methods and the reporting of data limitations, few possess the ability to be audited. Data within the last 3 to 5 years are available from all six databases. Service time, in total, and by component (e.g., pre-service, intra-service, and post-service) is most available for surgical services. NAMCS Office provides an estimate of intra-service time for E&M services; but pre- and post-service times are not available. Most appear to possess desirable statistical properties with respect to sample size and the ability to assess outliers; however, there is less certainty with respect to representativeness across the large number of stated desirable characteristics.

Of all the databases evaluated, the NSQIP may be best suited for determining service time for surgical services. Further evaluation of its representativeness and statistical properties would be a logical next step. However, our review of the above extant databases suggests that primary data collection will be required for most other types of services, especially E&M services and office-based procedures and diagnostic testing, unless there is major revamping of the survey data collection methods.

2. THE FEASIBILITY OF COLLECTING PRIMARY DATA

2.1 Initial Telephone Interviews with Large Health Systems

A primary objective of this study is to assess the feasibility of primary data collection that would provide time estimates from a cohort of physician practices. To do so, we started by holding telephone interviews with five large health care organizations to assess what time information they current collected. The purpose of the telephone interviews was to obtain preliminary answers to the following questions:

- Does the organization routinely collect the time that is required to provide a clinical service, specific phases of the clinical service, and, for what broad types of services? Are the time data linked directly to CPT codes?
- Are time data available for different types of providers and clinical staff that provide support services to clinicians?
- What are the time data sources?
- For what sites of service are time data available?

We reported on the results of these telephone interviews in a previous communication dated February 8, 2011, which is included as **Appendix B**. In summary, the five organizations with whom we held initial interviews have varying degrees of sophistication and experience with primary data collection of clinical service time. No organization has collected time for the purpose of physician payment, but all felt that primary data collection of clinical service time was feasible. Thus, it is highly unlikely that there exist data repositories that contain physician time data that would be appropriate and available for use as objective measures of time for updating the Medicare PFS. Rather, our initial interviews led to the conclusion that indepth interviews needed to be conducted to more fully understand the potential capabilities of electronic data systems and the method of direct observation for a fuller assessment of prospective primary data collection feasibility from a cohort of practices.

2.2 Site Visits: Structure and Planning

We selected two health care organizations (designated systems A and system B) for more intensive investigation. Both offer a full range of medical and surgical services to large populations and have reputations for their high quality of care. We arranged to meet with the leaders of these organizations at the policy and decision-making level and key persons with respect to conducting research and analysis of data. We prepared an outline for site visits to the two sites, entitled Site Visit Interview Protocol for Assessing Feasibility of Primary Data Collection, attached in **Appendix C** and developed agendas at the two sites based on that document. Key questions for the meetings were:

- Did the organization have the capacity to provide objective data on the time it took providers to perform medical services?

- Could time information be obtained on a wide variety of services- spanning medical, diagnostic, surgical, imaging and laboratory services?
- What methods could be used?
- What limitations with might be expected with respect to the reliability, validity and possible application of data that might be obtained?
- What support could be expected from the institutional leadership of each organization, physicians/other providers and research staff for investigations that might be conducted?

2.3 Summary of Findings from Site Visits at System A and System B

The findings of our explorations at these two sites are encouraging with respect to the major objectives of this project. We summarize the salient findings of our site visits in the text below and present more detailed information, formatted to follow the site visit protocol, as **Table 1** following the text.

- The top policy leadership and the principal researchers at both organizations voiced support and interest in meeting the objectives of the research and in working with others to achieve them.
- Both systems are broad-based, with large populations under care, physicians in all the major specialties as well as a broad range of non-physician providers, and well-organized research divisions. Both organizations are known for the high quality of the health care they provide and for their interest in furthering both the efficiency of that care and the satisfaction of physicians and other professionals who provide it. (**Appendix D Table 1: Types of Providers**).
- Both organizations have gathered extensive objective data on time. In the case of surgery and other procedures performed on an inpatient or ambulatory basis, information on time is recorded in OR logs. While there are differences in the methods used at the two organizations, those differences appear to be complementary in that they may provide information with respect to our major objectives from somewhat differing perspectives.
- Both organizations code all individual services using CPT and can link the times of specific services to the billing codes (**Appendix D Table 2: Data on Time for Selected Services at System A, Appendix D Table 3: Data on Time for Selected Services at System B**).
- System A has for the past ten years obtained information regarding office-based services primarily by direct observation, a process in which every staff member takes an active part at for at least one week a year. This approach, while labor intensive, has proven highly productive for this organization in making improvements in quality, efficiency and professional satisfaction in a large number of targeted investigations. Direct observation is but one component of what is a major investment in operations research. The organization patterned its approach on methods for continuous quality improvement. The organization is also involved in pioneering studies to make some of these observations electronically.
- System B, with similar objectives with regard to the quality of care and patient satisfaction, has also collected extensive information on the times required to provide high quality care in an efficient manner. This organization relies primarily on

exploiting its electronic health system and such administrative sources as schedules and appointment times to measure time objectively.

Table 1. Summary of Findings from Site Visits Systems A and B

	Part I	System A	System B
1	Has Organization Captured Time (in minutes)?	Yes	Yes
1a	Purpose of capturing time	<p>Improve efficiency, patient experience, and quality of care through dedicated system of continuous quality improvement</p> <p>Functional objective is “flow process” by physician and medical assistant (MA) and appropriate use of physician time. System A employs average durations of clinical services to determine staffing needs under conditions where demand for clinical services may vary over time. .</p>	<p>Increase efficiency, patient satisfaction and quality of care. Similar objectives to those of System A</p>
2	Most feasible source of time data	<p>Direct observation for most services</p> <p>Average times for scheduling selected frequent services</p>	<p>Electronic medical records (EPIC EHR) (System B’s proprietary information system) Appointment times Throughput</p>
3	What components of time do you collect?	<p>Finely granular breakdown of lead time, including pre, intra, post physician time, time of medical assistants, tech workers through entire patient experience . Well described in System A’s teaching materials</p>	<p>Map every computer click in primary care Map to encounters on system’s IS Calculate Time of Doctor Office Visit (DOV) monthly in each primary care region</p>
3a	Comment	<p>Includes direct observation within intra-service period</p>	<p>Measure time that physician is in examination room, but not direct observation</p>
4	For which types of providers can time be collected? (see Table A2 for few exceptions in specialties of importance to Medicare-)	<p>Full spectrum of specialties ~ 600 physicians</p> <p>There are several clinical centers</p> <p>No chiropractic services</p>	<p>System B has full spectrum of specialties</p> <p>Some are non-plan FFS patients. CPT coding of all services</p> <p>No chiropractic services</p>
5	Sites of service for time data	<p>Office, ambulatory surgical center, hospital, pathology lab, imaging, nuclear medicine</p>	<p>Office, ambulatory surgical center , hospital, imaging, nuclear medicine</p>

Table 1. Summary of Findings from Site Visits Systems A and B (cont.)

	Part I	System A	System B
6	Method or process used to collect time data	Primarily direct observation	Primarily using electronic data expert opinion. schedule times
6a	Do you develop data collection protocols?	Yes. Research Institute	Yes. Dedicated Research Division
6b	Are there individuals in the organization who are likely to provide leadership for objective measurement of time of clinical services?	Yes, CEO (MD) COO/CIO, Physician research leaders	Yes, CEO (MD), Research director, physician department heads
6c	What staff involved in collecting time data?	All Physicians (incl. CEO); administrators, research staff all must do annually	Research staff
6d	Frequency of data collection	Performed full time In one week cycles	Performed full time
6e	Sample size	Variable	Variable
7	Data limitations and caveats	Question extent of impact of efficiencies on physician time	
7a	Would organization share nature of analyses and key findings?	Yes, would consider collaborative research	Yes, would consider collaborative research
7b	Major strengths	Meticulous attention to data collection by direct observation with participation of entire organization from CEO on down. Full-bore commitment to operations research through proprietary system of continuous quality improvement	Quality of research team, acceptance of research and support by providers. Interest in potentiating information system (e.g. tracking EHR clicks to specific patient encounter)
7b	Limitations	Direct observation method is labor intensive.	Do not observe encounter in exam rooms
7c	Potential biases	Hawthorne effect thought to be minimal (direct observation in use >10 years)	Hawthorne effect thought to be minimal Potential errors using electronic data systems would need to be evaluated

Table 1. Summary of Findings from Site Visits Systems A and B (cont.)

	Part I	System A	System B
7d	Representativeness	Focus on flow and optimization of physician time (by offloading to 1:1 medical assistant) for appropriate work reduces physician time and increases MA time and other lower dollar rate inputs to practice expense. Probable systematic shift in resource inputs from physician time and work to practice expense (personnel subs, technology, re-engineering) This issue may apply primarily to primary care services	Same as at System A Physician time reduction achieved by increasing inputs of other personnel, technology, re-engineering (e.g., disease management) These tend to increase practice expense as physician time decreases
8	CPT coding, incl. quality	Done for all services for billing purposes, not for payment of individual physicians, who are salaried. Physicians select CPT code. Coding taught to providers, monitored statistically. EHR has code-check capability, but not used. Quality of coding is probably comparable to usual coding in FFS sector	Done for all services, but not for billing of subscribers. System has subset of FFS patients. Code-check system (based on documentation in EHR) used for coding, which is confirmed by physician. Quality of coding is probably superior to coding in FFS sector
8a	How could you conduct the linkage of time data to CPT ?	Merge separate data bases (direct observation, OR logs and CPT)	Merge separate data bases (electronic, OR logs and CPT)
8b	Confidence in accuracy of linkage	High. Not seen as a problem	High. Not seen as a problem
8c	What elements of time could be linked?	All, except time out of facility not on EHR (e.g., work at home off line, some phone time, "curb consults")	All, except time not spent on EHR, e.g. phone time not logged on EHR, hospital care time not on EHR, e.g. discussions w nurses, hosp. room encounters.

Table 1. Summary of Findings from Site Visits Systems A and B (cont.)

	Part II	System A	System B
1	Willingness to prospectively collect and share with MedPAC or CMS	Yes, strong support and interest at institutional level.	Yes, strong support and interest at institutional level
1a	Willingness to accept third-party participation in data collection.	Yes	Yes
1b	Willingness to follow prescribed protocols	Yes, based on agreement with protocols	Yes, based on agreement with protocols
2	Willingness of conduct direct observation studies	Yes. This is the basic process research methodology of organization	Yes, but probably limited. Not basic process research method, but interest in use in establishing quality of times based on EHR.
3	Cost estimates of data collection	Not currently estimable. Function of currently unknown scope, sample sizes, study designs. Infrastructure (Research institute, facilities, computer systems, software, experienced staff in place)	Not currently estimable. Function of currently unknown scope, sample sizes, study designs Infrastructure (Research division, facilities, computer systems, software, experienced staff in place)
4	Would comparative data from other participating practices be valuable to your organization?	Yes	Yes
5.	Thoughts/suggestions:re collecting objective time at CPT level	CPT already of use in evaluating for which services losses and gains to this organization occur. If it results in correction of distortions in physician payment it will benefit U.S. health care generally, e.g. through possible support of career choice for primary care and resource allocation nationally	Even though organization itself is salaried staff model, research division is interested in developing its own software for use in FFS sector, therefore strong interest in CPT structure. Strong support for doing this for national priorities, including consequences of improving the fee schedule, including allocation of resources across sectors and specialties, support of primary care activities.

3. DISCUSSION AND RECOMMENDATIONS FOR SPONSORED STUDIES

The findings from our site visits to two large health systems reveals that it is feasible to obtain objective measures of physician time using multiple methods of data collection. To do so will require carefully planned studies for the specific purpose of obtaining time that is appropriate for Medicare payment to physicians and non-physician providers across a broad spectrum of services and sites of service. It is our recommendation that pilot testing of a mixed-method primary data collection approach first—parallel studies by direct observation and electronic methods—occur in Systems A and B and then be applied more broadly across a larger set of practices to allow for an assessment of the feasibility of primary data collection across practices that have multiple specialties, including variety of specialists who perform E&M services. ~~that have organizational diversity~~. It may very well be that primary data collection is only feasible in clinical settings that have a robust electronic data systems and/or a culture of process improvement, characteristics of both Systems A and B. Such a finding may lead to the conclusion that it is not feasible to obtain objective time measurements that are broadly representative across all practicing providers but that it is feasible to obtain objective time measurements from practices that may be more efficient than average. Below, we provide primary data collection recommendations for seven sets of services.

The two major approaches to primary data collection on objective measures of time (in addition to operating room logs), direct observation and electronic methods, have strengths and weaknesses. Direct observation is labor-intensive, hence relatively costly, and possibly biased due to Hawthorne effects. However, considering the detail with which it can be performed, direct observation is inherently a more reliable and internally valid method for obtaining data on the times of the most common office-based E&M services (and the codes that account for the most Medicare dollars) than indirect methods employing electronic timestamps. On the other hand, direct observation at a large number and variety of practices is probably not feasible, given the labor-intensive nature of the method and the issue of intrusiveness at practices where it is not already part of the culture and readily accepted. The measurement of the time that providers devote to specific services has not been a major focus of developers of electronic medical records; such methods will need to be developed or refined. Such methods hold the potential for efficient data-gathering on a broad scale; however, their reliability and validity for measurement of service times must be established by careful studies. A chief purpose of the studies we propose would be to assess the strengths and weaknesses of these two approaches.

3.1 Office-based E&M Services and Hospital E&M Services

Given the importance of these services as a substantial fraction of overall physician payments, we recommend that parallel studies by direct observation and electronic methods

(e.g., time stamps by EHR) be performed at Systems A and B as a means of calibrating electronic methods. Once their reliability and possible systematic departure from the direct observational gold standard values are established, properly validated electronic methods could then be employed at a larger sample of clinical sites, including large integrated health care providers, large group practices, and possibly small primary care practices.

With respect to hospital E&M services, such as initial hospital care (admissions), subsequent care (both medical and post-operative care), and discharge services, direct observation should be performed, at a minimum for the purpose of establishing “gold standard” values against which less labor-intensive but indirect methods such as use of data from EHRs might be compared. The interruptions, discontinuous care, patient-related communications with physicians, other professionals and relatives, and multiple sites at which such care takes place (patient bedside, nursing station, x-ray department, etc...) leads us to believe that there is no ready substitute for direct observation in providing objective data on this arena of physician services.

Recommendation 1. Perform studies of office, hospital, and nursing facility E&M services by direct observation. Perform parallel studies using electronic methods. Conduct statistical analyses of the reliability and validity of each.

3.2 Common E&M Services that may be Performed Differently by Different Specialties, including Non-Physician Providers

While it is essential that valid measures of time be obtained for the primary care specialties, which perform the bulk of these services, it is important that these values not be automatically applied to physicians in other specialties who may provide quite different services under the same codes and levels of service. It will be important to determine whether post-operative surgical care (by a variety of important surgical subspecialties) and subsequent hospital care for medical conditions are provided with times that are the same or different. Discharge day services should be measured as well. The same should be done for consultations, where there may be systematic differences by specialty.

Office E&M services of important specialties, including cardiology, gastroenterology, neurology, gastroenterology, allergy and immunology, obstetrics and gynecology, dermatology, and ophthalmology (which has its own specialty-specific set of codes) should be studied, both by direct observation and electronic methods where this is feasible (in order to “calibrate” the electronic methods) and by electronic methods at a broader variety of sites.

Recommendation 2. Perform studies of office E&M services by direct observation in a set of specialties, including non-physician providers, selected to reflect possible diversity in the content and time of services. Perform parallel studies using electronic methods. Conduct statistical analyses of the reliability and validity of each.

3.3 Office-based Surgical Procedures, Imaging Services and Professional Component of Diagnostic Services, including Pathology Services

Direct observation is inherently a more valid and reliable method of measuring the time of non-E&M services that take place within an office-based setting, such as dermatologic procedures and office-based endoscopy, as these services generally take place within the context of an office visit that includes multiple major activities.

Further, there may be no good substitute for direct observation services in which physicians interpret diagnostic tests performed by others (e.g. audiometry, spirometry, etc.), or laboratory studies performed by pathologists. The times for physician interpretation for certain types of tests, such as electrocardiograms, cardiac ultrasound, long-term electrocardiography (Holter monitoring), electroencephalograms, or other tests read in batch form at many institutions, may initially be estimated by examining physicians' schedules for such work (i.e. throughput); however, it is likely that direct observation will be required to validate per service time. A more detailed discussion of this type of approach may be found in the June 2011 MedPAC Report to Congress (Chapter 1, pages 19-20).

With respect to the interpretation of imaging procedures, direct observation (and possibly the use of office schedules) combined with billing data, will bring greater validity to the time values for these services. Information from organizations that perform tele-radiography services (which provide interpretation services exclusively) may be particularly useful in understanding the times required for radiology interpretation.

Recommendation 3. Perform studies by direct observation of office-based invasive procedures, such as office-based endoscopy, that co-occur with other major office-based services and of physicians' interpretations of diagnostic studies performed by others or of laboratory studies performed by pathologists. Further, perform studies of direct observation and the use of times derived from office schedules or billing data to generate valid estimates of time required to perform radiology interpretation services and other diagnostic tests read in batch form.

3.4 Surgical and Diagnostic Services Performed in Ambulatory Care Facilities

With respect to surgical procedures and diagnostic procedures (e.g. colonoscopies) commonly performed in ambulatory facilities, operating or procedure room logs can be used for the acquisition of data on physician time.

Recommendation 4. Obtain information on the time of surgical and diagnostic procedures most frequently performed in ambulatory facilities from operating or procedure room logs.

3.5 Major Surgical Procedures

Operating room logs can provide reliable time data on essentially all the major surgical procedures. As with other services, it will be important to link these data with the corresponding CPT codes. There should be no difficulty in obtaining valid times for all major surgical procedures in this manner.

Recommendation 5. Obtain information on the time of surgical procedures most frequently performed in hospitals from operating room logs.

REFERENCES

Annual Report to Congress. Washington, DC. Physician Payment Review Commission, 1988, pp 71-82.

Annual Report to Congress. Washington, DC. Physician Payment Review Commission, 1989.

Braun, P., Hsiao, W.C., Becker, E.R., DeNicola, M. Evaluation and management services in the resource-based relative value scale. *JAMA* 260:2409-17, 1988.

Braun, P., Dernburg, J., Dunn, D.L., Cohen, W. Predicting the work of evaluation and management services. *Medical Care* 30:NS61-NS79, 1992

Hsiao, W., Braun, P., Yntema, D., Becker, E.R. Estimating physicians' work for a resource-based relative value scale. *N Engl J Med* 319:835-841, 1988.

Leape, L., Freshour, M., Yntema, D., Hsiao, W. Small-group judgment methods for determining resource-base relative values. *Medical Care* 30:NS 28-39, 1992

McCall, N., Cromwell, J., Braun, P. Validation of physician survey estimates of surgical time using operating room logs. *Medical Care Research and Review*. 63: No. 6, 1-14, 2006

Medicare and the Health Care Delivery System: Report to Congress. Washington, DC. Medicare Payment Advisory Commission, June 2011.