

March 28, 2018

Francis J. Crosson, MD
Chairman
The Medicare Payment Advisory
Commission
425 I Street, NW, Suite 701
Washington, DC 20001

Dear Dr. Crosson:

We are writing to respectfully but firmly request that the Medicare Payment Advisory Commission (MedPAC) remove any references to the University of Minnesota contractor study (Zismer et al, 2014/2015)¹ from the Commission's June 2018 Report's chapter on "Rebalancing" the Medicare Physician Fee Schedule (PFS). The American Medical Association (AMA) has serious concerns about this entire chapter, which makes repeated assertions that are unsupported by any data and ignores factual information provided by the AMA and the American College of Surgeons that contradicts or mitigates many of these assertions. The inclusion of references to the Zismer study is only one of the problems with the chapter; we find the use of this study as a justification for dramatic redistribution of Medicare payments to be particularly troublesome given the small number of participating practices, inconsistencies and shortcomings in its data, and destruction of the code level data that would have been necessary for any outside validation of the study findings. It is highly ironic that the draft MedPAC chapter complains about the "objectivity and quality of the data" that underpins the current fee schedule values and then wants to replace these values based on results of a study of four specialties and five practices that even MedPAC staff has conceded is "not definitive."

Our concerns are based upon our recent electronic mail and telephone dialogues with several MedPAC staff members,² and upon our comprehensive review of related Commission documents. At this time we also wish to gratefully acknowledge the staff's accessibility and their forthcoming responses to our queries.

I. Context

The AMA agrees with MedPAC that reliable physician service time methods and measurements are vital elements required for the proper functioning of the Medicare Resource-Based Relative Value Scale (RBRVS) that underpins PFS payments. Validating PFS service times using a "top down" approach (as discussed at MedPAC's October 10, 2014 meeting and more fully described in a 2012 Zismer report) could potentially be used as another tool to verify the accuracy of cross-specialty relativity within the

¹ Zismer DK, Christianson J, Marr T, and Cummings D. An Examination of the Professional Services Productivity for Physicians and Licensed, Advance Practice Professionals Across Six Specialties in Independent and Integrated Clinical Practice. University of Minnesota School of Public Health: MedPAC Contractor Report, submitted July 2014. An updated version was submitted in July of 2015.

² Mathews JE, PhD, Executive Director; Hayes K, PhD, Consultant; and Winter A, MPP, Principal Policy Analyst.

RBRVS if several issues with collection of adequate and representative data can be resolved. However, the AMA strongly contests any direct or implied assertion that the later Zismer report (variously referred to as either the 2014 or 2015 report) accurately demonstrates large and pervasive PFS time discrepancies that can be generalized across all Medicare specialties and providers.

Results from this study are inadequate to justify across-the-board payment redistributions such as those envisioned in MedPAC staff presentations at the January 2018 meeting and the forthcoming June report chapter on “Rebalancing Medicare’s Physician Fee Schedule.” While the exact size and shape of this redistribution is unclear and has morphed since the Commission’s discussion in January, it apparently could entail a 10 percent increase in ambulatory Evaluation & Management (E&M) funded by corresponding reductions in payments for all other PFS services. Depending on how “ambulatory evaluation and management” is eventually defined, reductions in payment for other services—possibly even including recently-introduced care coordination and management codes expected to benefit primary care physicians—could range from 3.8 to 4.5 percent.

As noted earlier, in a blog about the study, MedPAC staff acknowledged that its results are “not definitive” and that there were “challenges” in “collecting enough data to reliably conduct this analysis.” They chose to use these “provisional findings” for purposes well beyond what the study’s authors intended, however, because the findings were “directionally consistent with the Commission’s concerns—that primary care services are undervalued in the fee schedule and procedural and testing services are overstated.” Another relevant study conclusion questioning the “generally accepted conclusion of a shortage of primary care physicians in the US” has been completely ignored.

In an attempt to understand the Zismer findings, the AMA reviewed the Zismer studies as well as the January 2018 and October 2014 meeting transcripts. We then sought additional clarification from MedPAC staff and learned that at staff’s direction, the researchers “did not retain” these data. Based on what was available, however, the AMA finds no reliable evidence of time discrepancies that would justify the proposals that have been presented to the Commissioners and modeled in the draft chapter. Instead, we have identified several troubling issues that are listed below and then addressed in more detail, using illustrative material from study-related documents.

II. Key Concerns

- 1. This Zismer report (variously referred to in MedPAC documents as a 2014 or 2015 study) and a 2012 predecessor were aimed at testing a concept that could be used to identify services where RBRVS times may be inaccurate. They are not, and did not purport to be, a representative sample of practices. Neither report called for payment changes prior to further analysis.**
- 2. The analyses conducted of actual hours worked (based upon group practice-specific block schedules) compared to projected hours worked (based on simulated RBRVS hours worked) suffered from design, data collection, and interpretation flaws. The methodologic rigor of this analysis falls considerably below that normally associated with MedPAC’s work.**
- 3. The generalizability of the Zismer 2014/2015 study findings also was not rigorously addressed, but the size, scope, location and study design all cast serious doubts about**

whether the findings from this limited study can be extrapolated to all Medicare specialties and providers and used to set national policy.

- 4. The report on the 2014/2015 Zisner report's conclusions regarding differences between actual times and RBRVS times does not include a transparent discussion of how the comparisons between times were calculated or any code level data that could be used to verify the study's conclusions, methods and results. In fact, to ensure confidentiality to the participating practices, all data that could be used to verify the results have apparently been destroyed.**

III. Discussion in Detail

- *The focus of the Zisner studies 2015 study was on efficient physician services delivery patterns, the degree of overlap between work of physicians and advanced clinical practitioners, and the feasibility of using billing and scheduling data from physician practices to verify time used by the RBRVS. Attention was targeted on description of patterns,³ as used by "larger subspecialized" and "well-organized group practices."⁴*
 - The authors repeatedly characterize their work as exploratory rather than definitive. Illustrative examples include:
 - "The purpose of this study was to provide information and insights on how this topic can/should be examined on a larger scale;"
 - "To provide MedPAC officials with perspectives on a limited sample of specialized medical practices;"
 - "...provide the foundation for a model for future, more in-depth examination;"
 - Follow-on studies should include "development of methods to estimate, with precision, physician allocations of time;" and
 - "In gaining such understandings through this proposed pilot project work to devise a reliable and efficient model (and method) to evaluate such efforts in an expanded sample of specialty practices across a range of U.S. healthcare system designs."
- *The analysis conducted of actual hours worked (based upon group practice-specific block schedules) compared to projected hours worked (based on simulated RBRVS hours) suffered from design, data collection, and interpretation flaws. The AMA has multiple concerns about the rigor of this study and the extent to which the methodology and findings were documented. The methodology used and analysis employed were intended to assess work allocation patterns among practitioners. Details presented about service time measurement methodology and results analysis are incomplete, including the absence of basic information normally included in reports of this kind (e.g., clear and consistent information on participant numbers, composition and characteristics).*

³ The most common patterns used by these practices involve predetermined allocation of work between physicians and their associated advanced practice clinicians.

⁴ Productivity as measured by work relative value units

- *The actual numbers of participants in the secondary investigation overall and in each specialty analyzed are unclear.* For five specialties/six group practices, the numbers can be deduced using histogram bar labels in Exhibits C and D for each specialty along with Appendix E. However, even after adjusting for the fact that oncology is included in Exhibits C and D but not Appendix E, there are unexplained inconsistencies. For example, Appendix E shows one more urology advanced practice clinician (APC), one more orthopedic surgeon and five more orthopedic APCs than can be found in Exhibits C and D. In addition, interview summaries in Appendix D indicate that the orthopedic practice actually had more than 30 mid-level practitioners. For radiology, the participant number is unavailable even through inference. It also appears by inference that some practices chose a physician subset to participate (i.e., data are shown for just one oncologist and four family medicine physicians while all participating practices were termed “large” organizations). Oncology apparently was not included in the final results but family medicine was. Given that both oncology and family medicine include very small samples, why was one included while the other was excluded? Physician subset selection criteria are not discussed.
- *Data collection periods were not standardized.* While the study design called for a period of one year, the actual observation period varied across practices, from less than one month up to a full year. The rationale for and the potential impact of this variation on study results are not addressed (e.g., decreased reliability and generalizability of conclusions).
- *Criteria for judging data adequacy for analysis were neither provided prospectively nor explained post hoc.* It is evident from Radiology Exhibit A that data were collected for radiologists yet their data were described as “inadequate” and no Appendix E graphs were included for them. Possibly this is because this practice had not incorporated APCs but times for the physicians in the practice presumably would have been available. Similarly, Oncology Exhibits A-K demonstrate data collection from medical oncology physicians and APC providers, yet there are no Oncology graphics in Appendix E and no explanation is offered for that absence. “Data limitations” reportedly precluded calculation of the average number of units and work Relative Value Units (RVUs) per service category in the family medicine practice but there is no explanation as to what these limitations are or how they might have affected estimated times per service.
- *The overall precision and validity of the actual hours worked data are unclear.* From Appendix A, it appears that actual hours worked were extracted by researchers from clinician schedule block assignments that were obtained from group practice managers. No narrative description or any example is shown of a schedule assignment document nor is there evidence that all managers used the same or similar document. For example, each practice determined which categories of services to report and which Current Procedural Terminology® (CPT) codes to assign to which category, which could have created related variation in the study results, especially for codes which could have been reported in more than one category. A blank Data Display Model Category #1 is provided under “Approach to Analytics (for study work team),” but no data table or any example of a data display model document as completed by a researcher is included. The hours data are referred to as “estimated physician work effort (time)” (emphasis added). No mechanism is described for validation by the researchers that the practice managers’ schedules in fact reflect actual clinician hours worked (e.g., by cross-check with Operating Room records for procedural time blocks). No discussion is provided about

the potential impact on study results of non-standardized practice manager inputs (e.g., decreased reliability and generalizability of conclusions).

- *It is unclear if and how exceptions to the actual hours worked attribution process would be identified and handled by practice managers and researchers.* The AMA is skeptical that clinically active clinicians adhered completely every day to block-diagram activity schedules with hard start and stop times created by practice managers. When a clinician began a block earlier than its scheduled start or ran over its scheduled end, how was the additional time worked captured for the study? If a clinician used a scheduled break to go to another site to see a patient she/he regularly follows, or to respond to a call for an urgent subspecialty consultation, how was that work time detected and recorded? How were paperwork demands such as certification, documentation and preauthorization requirements recognized in the study? If a clinician used time outside of the block (and perhaps off-site and off-hours) to perform these tasks or complete electronic medical record documentation or to review patient results, how would practice managers or researchers be made aware? Do the schedulers keep track of time that physicians spend doing patient rounds at the hospital before and after they come to the office to see patients? When unplanned schedule changes are imposed upon a physician by others, how are they handled? For example, is there a documentation of cases such as a surgeon whose scheduled elective procedure gets bumped by another surgeon's emergency case? The surgeon may not be able to readily fill this time with other billable activities, as described by practice managers to the researchers as the expectation for the use of down time. Is this treated as a "break" that therefore falls outside the study's definition of actual hours worked? All of these scenarios are much more easily imagined than lock-step adherence by all study clinicians to highly-regimented block diagram schedules devised by non-clinicians. What efforts were made by researchers to detect such events, such as interviewing one or more study clinicians from each practice (other than the study's "practice representative")?
- *Solutions used for two significant potential ambiguities in attribution of hours worked raise concerns about accuracy and validity of the study results and their subsequent interpretations.* Rationales for the solutions chosen are not provided.
 - "Incident to" services. The time for services performed by APCs whether billed as incident to or independently-provided (i.e., with or without physician supervision) were attributed to the APCs. No time from incident to services is attributed to the supervising physicians, presumably reducing estimates of their actual hours worked while inflating the hours worked by the APCs. This decision offers a likely explanation for why the researchers concluded that physicians spent less time and APCs spent more time than would have been indicated in the RBRVS times. The choice to bill APC-performed services, and thereby attribute actual and simulated hours worked, as incident to or independent, was not subject to any guidelines from the researchers, introducing yet more variability into the results of the study.
 - Global Surgical Services. A global surgical service includes pre-defined E&M encounters that occurred either pre-operatively, post-operatively or both. Since 71 of the 76 participating physicians (93 percent) come from procedurally-oriented specialties (all but four family physicians and one oncologist), inaccurate attribution of global surgical service package hours could have an

enormous impact on the results of this study. But it is impossible to determine exactly how the researchers dealt with this question. For example: How actual procedure-performance time was parsed out and attributed to various physicians is unclear. The “Approach to the Study (Methods)” section states that “...the surgical procedure was reported with its CPT code” (emphasis added). This suggests that the entire RBRVS package time was attributed to the operating physician on the date of the procedure, an attribution that will inflate that physician’s simulated RBRVS hours worked on that date without changing the actual hours extracted from practice manager schedules.

- The study goes on to state that “A pre-operative or post-operative visit...in a procedure’s global surgical package was reported with its CPT code” (emphasis added) Which CPT code is this referring to:
 - One possibility is that the study is referring to a modifier that can be used to identify pre-operative services and/or a code that can be used to identify a follow-up visit. Because they are part of the global surgical fee, however, these services would not generally be paid for, so in many practices, they are not separately reported. Would all the practices in the study have actually known when someone other than the principle surgeon performed a pre-procedure service or post-operative visit? Also, since there are no times for these codes in the RUC data base, what times would the researchers have used in their simulation of RBRVS times? Only when both the procedure and all of the global package post-operative visits are furnished by the operating physician will all of the simulated RBRVS hours worked be correctly attributed, an uncommon situation when busy practitioners rotate work sites weekly (as study physicians were said to do).
 - Or does “its CPT code” again refer to the global surgical service CPT code, so that the global code was used to track each associated post-operative visit, wrongly attributing significant simulated RBRVS hours worked to any physician furnishing a post-operative visit?
- *Study results have not been, and cannot be, subjected to the kind of outside scrutiny that should be a prerequisite for policy changes of the scale that MedPAC is contemplating.*
- *The actual and simulated RBRVS hours in Appendix E are limited and not easily interpreted.* Results are provided for only five practices/four specialties without stated inclusion/exclusion criteria. Unfortunately all 10 results graphs are presented as calculated ratios—percentages by which “CMS Estimate” (simulated RBRVS) hours worked diverge from “Practice Estimate” (practice manager block-diagram based) hours worked—rather than as actual hours data. Determining actual hours or their proxy Practice Estimate hours from the provided graphical results is impossible, since no data tables or any example linking hours to the calculated ratios are provided.
- *Publicly presented study results are not transparent and cannot be validated.* Clearly the Practice Estimate hours worked raw data were available to the researchers at some point for constructing the Appendix E graphic results. Even more importantly, the raw data were used to create a results graphic shown as Slide 5 during the January 12, 2018 MedPAC meeting and with the Zismer 2014 study cited as the source. The y-axis of that

graph is labeled “average hours per day” and histogram bar sets are shown for “Fee schedule time” and “Hours worked” for four specialties (Cardiology, Family Medicine, Orthopedics, and Urology). Neither the hours data nor the graph appear in the Zismer 2014 report, however.

- In response to AMA queries about the data behind the results graph shown to the Commissioners, MedPAC staff confirmed to AMA that there is no additional appendix or technical supplement beyond the Zismer 2015 report posted on MedPAC’s website. Staff also stated that MedPAC contractually prevented the researchers from “retaining any data associated with this work.”
 - *Despite the previously-conceded study caveats and the disposal of source data for examination and validation of the contractor’s findings, the study results were confidently presented to the Commissioners without any mention of study limitations. Nor was there any reference to the October 2014 Commission discussion where various issues with the study were raised. While viewing Slide 5, Commissioners were told that “This slide illustrates the extent to which certain services have become overpriced” and “The contractor’s study found that the hours assumed in the fee schedule exceeded actual hours worked for physicians in all four practices”.⁵ Only eight Commissioners of 16 present for the January 2018 presentation had also been in attendance at the October 2014 presentation, suggesting that a number of members were unacquainted with its limitations.*
- **The generalizability of the Zismer 2014 study findings also was not rigorously addressed, casting serious doubts about whether the findings from this limited study can be extrapolated to all Medicare specialties and providers, and can be used to set national policy.**
 - The total number of physicians for whom at least some time data were reported and analyzed is not clearly stated in the report but by AMA’s calculation (described above) appears to be 76. This group is comprised of six group practices/five specialties, and 93 percent (71/76) of the group members are in procedural specialties. It is difficult to imagine that group constitutes a representative sample of the more than 500,000 physicians enrolled in the Medicare program.
 - The locations of the participating group practices are not explicitly stated but references to hospitals in which study physicians practiced indicate that all were in the Minneapolis-St. Paul metropolitan area.⁶ This does not represent the geographic diversity involved in the practice of medicine. As noted in the paper, state scope of practice laws could affect the way work and time is allocated between physicians. Patient mix, teaching status and a variety of other factors could also play a role.

⁵ MedPAC. Public Meeting Transcript. Washington, DC., January 12, 2018, pp 59-60.

⁶ This is consistent with the Zismer 2012 report in which all the study practice locations are clearly identified as located in the Minneapolis-St. Paul metropolitan area.

- The use of simulated RBRVS physician work times as a comparison set for the study estimated actual hours worked is flawed. The procedure times provided in the RUC database are intended to represent the typical patient – no information is provided in the Zismer study about the characteristics of the patients provided services. The study does not tell us anything about the patient mix for the practice participants. We do know that both private pay and Medicare patients were included, however, so it is likely that some practices in the study would have had a patient mix that included many younger, healthier patients whose care might have been less intensive and time-consuming than typical. It is also important to remember that Minnesota patients in general are healthier and less likely to need very intensive care than patients in many other parts of the country,
- It is unclear from the two contractor reports whether some group practices participated in only the 2012 study, the 2014/2015 study, or both. For two specialties, the charts indicate that the services were performed in March of 2013 but dates are not provided for the other specialties. Notably, cardiologists numbered 44 in the 2012 study (explicitly stated) compared to 43 cardiologists (deduced by inference) in the 2014 study. Were these in fact the same group practice members? Were 2014/2015 report data for overlapping participants simply their 2012 data carried forward for repeated use?
- It is noteworthy that 44 Family Medicine physicians were tracked in the 2012 study but only four participated in 2014. The AMA is concerned about this dramatic decline in participants from this specialty between the studies since Slide 5 appears to show that physician service times are significantly less inflated for Family Medicine physicians versus the remaining “procedure-oriented” specialties. This observation then led to the policy proposal for increasing PFS payments for primary care services. The Zismer 2014/2015 report is silent on this question.

IV. Summary

For the reasons discussed above, the AMA concludes that this study does not provide sufficient or compelling evidence about potential Physician Fee Schedule time inaccuracies to serve as the foundation for major physician payment policy revisions. While we remain open to the concept of using practice data to help verify times in the RBRVS, we also do not believe that the work to date demonstrates that this is currently possible. At the very least, one would need to know that results were based on a large and representative sample of practices, not just four specialties and a handful of practices that are located in one of the healthiest cities in the country and that are big enough and specialized enough to make what the researchers regard as an “efficient” use of non-MD practitioners. The AMA, therefore, strongly recommends that MedPAC exclude the Zismer 2014/2015 report from use in the Commission’s June 2018 Report to the Congress chapter about Medicare PFS rebalancing.

We also ask that before PFS rebalancing is discussed any further by MedPAC, appropriate context is provided. In addition to including a discussion of the original purpose and limitations of the Zismer report, this would include a review of actual data regarding changes in Medicare payment levels and overall spending share for “ambulatory E&M codes” versus spending in other categories such as imaging and surgery. To assert that evaluation and management services have been passively devalued without ever calculating and reporting the actual changes that have occurred, creates the impression that MedPAC is afraid the data will not back up the Commission’s long-held views on the distribution of Medicare payments. The AMA has provided these types of data to MedPAC staff on several occasions and would be willing to do so again. However, MedPAC staff could certainly do these calculations in-house if they

Francis J. Crosson, MD

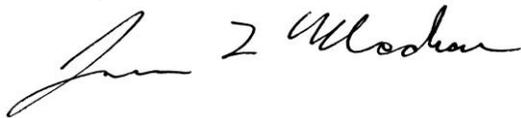
March 28, 2018

Page 9

do not trust AMA numbers. The point is that it is hard to “rebalance” something without knowing what the weights of the various components were in the first place.

We again thank Commission staff members for their accessibility to AMA staff for queries and for their forthcoming replies. We look forward to your attention to the issues that we have raised, and we are available to answer any questions that you may have.

Sincerely,

A handwritten signature in black ink, appearing to read "James L. Madara". The signature is fluid and cursive, with a large initial "J" and a distinct "M".

James L. Madara, MD

cc: Jon Christianson, PhD

James Mathews, PhD