Timothy Smith
Mark O. Dietrich

Second Edition

Excerpt from:

BVR/AHLA Guide to

Valuing Physician Compensation and Healthcare Service Arrangements

BVR
What It's Worth

AHLA
American Health
Lawyers Association

Chapter 26. On the Use and Misuse of Survey Data: An Interview With MGMA

CONTENTS

Chapter 26. On the Use and Misuse of Survey Data: An Interview With MGMA

Editor's note: Timothy Smith, CPA/ABV, senior managing director, Ankura Consulting, interviewed Meghan Wong, assistant director, data solutions, Medical Group Management Association.

1.0 Understanding What Survey Data Do and Do Not Represent

TS: *Meghan, first of all I'd like to thank you for taking time to do this interview for the* BVR/AHLA Guide to Valuing Physician Compensation and Healthcare Service Arrangements. *I should note for the readers that this is not the first time you and I have spoken about the subject of using survey data to set physician compensation.*

MW: Yes, we did a webinar for BVR back in December 2013 titled "The Use and Misuse of MGMA Data in Healthcare Valuations." We discussed various MGMA survey products and talked about how the valuation community was using MGMA's data. You also did a webinar for AHLA with another MGMA staff, Rachel Weber, in August of 2016 about the use of MGMA's data for compensation valuation. So, yes, we've talked about the use of MGMA's data by the valuation community going back several years.

TS: *Let's start out by discussing MGMA's interest in commenting on how its data are used in setting physician compensation in the marketplace, including how they are used in determining fair market value for regulatory compliance purposes. What's important for MGMA about how its data are used in the industry?*

MW: We want users to understand the credibility of the data we publish—all survey information goes through a rigorous editing process to ensure its accuracy. We also stress that MGMA data should be used to help guide decisions in conjunction with other information and not as a stand-alone metric. We also are clear that MGMA data should never be used to limit competition or be used in any way that is in violation of antitrust laws.

TS: *Many in the healthcare industry cite MGMA's data as representing the U.S. physician marketplace. I see stories in the trade press that say physician compensation has increased by X% because that was the increase for a given specialty in the most recent MGMA survey. In addition, lots of industry players frequently cite MGMA as indicative of national compensation levels for physicians. What's your reaction to these kinds of claims about the MGMA data?*

MW: We are not saying this is a one-to-one representation of the universe of medical practices that are in the country. This is strictly a survey sample. We do have the most comprehensive data set within the industry,

with over 80,000 providers in both physician-owned and hospital-owned practices, but it lends itself to being misquoted. There have been times in news articles where reporters will look to MGMA survey respondent demographics and say this is what the ownership trends look like for the country. We want to make sure we are very upfront about the fact that, while this is a robust data set, it is based on a sample of the practices that are responding to our survey.

There is the legal notice and disclaimer that we make sure we put in all of our surveys that states MGMA does not purport to offer advice that may be construed as specifically applicable to individual situations. The measures in our reports are descriptive statistics that are calculated from member and nonmember survey responses. We are really just providing this as a tool for consumers to utilize and inform them when they are building their own compensation plans or analyzing their practices. We definitely don't want any of our information to say that this is what it has to be or limit productivity or compensation in any way.

TS: *Would you explain what you mean by descriptive statistics, and how does descriptive statistics contrast with inferential statistics? As I understand it, for inferential statistics, random sampling or other techniques are used to develop a sample of a population that is intended to be representative of that population. An example of inferential statistics are political polls, where pollsters select a sample of voters and attempt to extrapolate to the entire U.S. electorate based on the sample.*

MW: MGMA survey data take a sample of the United States healthcare business population and report back descriptive statistics on that sample: mean, median, standard deviation, and percentile representations. Our data are meant to be utilized as a business tool and not a statistically significant data set that relies on more academic hypothesis testing and random sampling

TS: *So, when we think about and use industry survey data, we need to understand that the data are not designed to be statistically representative of the U.S. marketplace for physicians. Rather, the survey data are an accurate data portrayal or description of those practices and physicians who responded to the survey in a given year. Is that a correct understanding of what the data are and what they represent?*

MW: That is correct. Speaking for MGMA, our data reports are valid business tools and not meant to be used as academic statistical data sets, as is the case with other survey reports in the industry.

TS: *I'd like to explore the implications of knowing industry survey data reflect descriptive statistics rather than inferential statistics. Would this mean, for example, that one could not use the percentage of practices or physicians from health system-owned practice reporting to be the percentage of physicians nationally who work in health system-owned practices? I see presentations where individuals use this percentage in industry surveys such as MGMA as proof the majority of physicians in the U.S. are now employed in health system-owned practices.*

MW: Yes, we see various parties using the data in this way, even though the data are not intended to be used as an academic data set for extrapolating to the U.S. population of physicians. We try to inform such users about the actual nature of the data. For example, a *New York Times* article in the spring of 2010 misquoted MGMA Physician Compensation Survey demographic data as being predictive of the national trend. More hospital-owned practices had begun filling out MGMA surveys in previous years, yet the reporter extrapolated that to the entire country. Even though we reached out to correct that statement at the time of publication, that same "fact" comes back to us every year, with people wanting "updated data on the trend." MGMA continues to educate, but search engines on the internet have made the original article easy to access, which leads to continued misinterpretation of our data.

TS: *What about claims that certain physicians are among the most highly paid in the U.S. because their total compensation exceeds the MGMA 90th percentile? Is that a factual inference one can make based on benchmarking against the survey?*

MW: Of the sample of MGMA survey participants, those physicians that fall at and above the 90th percentile are among the highest providers in that data set. Again, as MGMA data represent descriptive statistics, these physicians at the 90th are not the "ceiling" of the country but rather the highest paid physicians in our particular sample.

Of note, there are outlier compensation figures above the 90th and below the 10th percentiles that are trimmed out of our data sets based on various review procedures we complete. First, we apply a statistical formula across all data sets that trims out (high and low) outliers, if present. Next, we manually look at the data for anomalies and large shifts in median from the previous year to identify whether perhaps only one group of physicians are skewing the data set up or down. Some of these trimmed figures may have been entered as the result of survey respondent error, but there are also cases where those outlier compensations are genuine for those individual physicians. Regardless of why the data were entered into our survey, singular outliers are suppressed from the data set to reinforce the delivery of a sound business tool. Our data analysis procedures support the goal of providing benchmarks that are not skewed by these outliers.

TS: *Let's talk about the regional data. MGMA provides descriptive statistics for its sample based on geographic areas or section of the country, as do many other industry surveys. What is the intention of reporting the data according to these sections and how did MGMA select the states that go into each geographic section?*

MW: MGMA identified the states that go into the four regions of Eastern, Midwest, Southern, and Western several decades ago. We have maintained these same regions over time to maintain the integrity of trending information across all of our survey products. MGMA does offer other regional breakouts as well, such as Minor Region and State.

TS: *Industry users often look to data from their geographic section as being a better indication of the marketplace for their physicians. Does MGMA observe that there are trends in each region that are unique to the markets in that region? Also, do you observe that regional statistics reflect trends in every state or market within the region?*

MW: The more detailed the data set is around a specific comparison practice, the more applicable it will be. Again, we urge data users to look to MGMA data as a guide and use discretion when applying benchmarks. If a data user is able to drill down to the state level, that is preferable to looking at just region. For example, comparing Colorado to the Western section is helpful, but it's even better to compare to just the state of Colorado so that other states are not mixed into the comparison (Utah, California, Nevada, etc.) For some markets, additional analysis of other data sources may be needed to supplement analysis of that area, for example, in states as large as California, where there are various differing markets even within the state.

2.0 The Correct Understanding and Use of Medians

TS: *Let's spend a little bit of time talking about the correct view of medians in survey data and how they should be viewed or understood by industry data users. Would you step us through how medians are calculated and what they mean or signify for a given data set?*

MW: We suggest that data users stay away from utilizing the mean as a representative of the survey sample. We urge people to instead use the median (with very few exceptions) as the representative of the sample because it is not going to be influenced by extreme outliers.

TS: *Is the median the most significant data point in a data set? Is the median the most frequently observed data point?*

MW: Median is where 50% of respondents fall above and 50% fall below in the sample. It is not the floor of the data set; rather, it's the central point. The median is a good starting point when looking at a metric, but data users should be comfortable using data that fall above and below that percentile, commensurate with practice characteristics, physician experience, and productivity.

Also, the most frequently occurring data point in a data set is going to be the mode, not the median. The mode can occur anywhere along the data array—it doesn't necessarily have to be the low, middle, or high point.

TS: *Many market participants in the healthcare industry, including a lot of valuators, think the MGMA median compensation per wRVU rate or the median total compensation level is the industry norm or national benchmark for what any given physician should make. How does MGMA think about these kinds of ideas about the reported median?*

MW: I think the median is a great high-level look at what a specialty is reporting that it earns across the nation, specific to our sample. But there are going to be many environmental factors or marketplace factors that play into what physicians are being compensated. It's also worth keeping in mind the total compensation figure is going to include other things external to productivity, such as on-call or medical directorship stipends. Total compensation by no means should be used as a base salary figure because it does include multiple facets.

When we say to look at the median as a representative, it is really just a starting point. Someone who is working in California is going to be earning at a median something much different than someone who is working in West Virginia. There are different marketplace factors that are in play. That is why we make MGMA data tools so robust in that you can filter by those different geographic sections, or ownership types, because those aspects affect compensation. The overall median might not be the catchall, be-all for all individuals out there.

TS: *To connect the issue of the median back to our prior discussion about descriptive nature of the survey data, can one claim MGMA medians are the national norms for physician compensation? Can one say the MGMA median is U.S. national median compensation level?*

MW: I think there is not one flat target rate for everyone within the marketplace. Compensation per work RVUs is no different than just regular compensation. There are going to be different factors at play, not just the location of the practice, but also what the provider's patient mix looks like and also the ability of the practice to reimburse on those procedures. There are many different factors that are involved with compensating particular physicians, and there should never really be a catchall for everyone regardless.

TS: *A lot people think the MGMA median total compensation amount is the amount of compensation needed to recruit or employ a new, but experienced, physician. They say it will take at least the median compensation level to incent an experienced physician to come to a new job in a physician practice. Does MGMA have any data or information that would speak to this belief about recruiting a new physician?*

MW: In actuality, we have a starting salary survey, and that is going to be more appropriate because it includes guaranteed compensation for new physicians. We found that new physicians are often earning less than

established physicians in a practice, as they haven't had time to build their panel size yet. All of the data within the starting salary data set are guaranteed salary figures for those first-year physicians in a practice, which is much more appropriate to use over the total compensation figure for established physicians. Also, within that data set, we allow the user to filter by whether the provider is new from residency or fellowship or whether the provider is experienced and just new to a practice.

3.0 Understanding Reported Percentiles as Part of a Data Set

TS: *What exactly are percentiles, and how are the data organized into percentiles?*

MW: We take a group of data points for a specific specialty and metric and report the corresponding position of the data relative to the group. So when we say a family medicine work RVU data point is at the 25th percentile, it means that 25% of those work RVUs falls below it and 75% of the work RVUs falls above it.

All data are arranged in order, from lowest to highest in value on the scale from 0 to100. Essentially, a straight line is drawn from one value to the next, and those lines and values are how we are able to provide descriptive percentiles for the values.

From a methodology standpoint, we don't report any values above the 90th or below the 10th percentiles.

TS: *Why do industry surveys report the specific percentiles of the 10th, 25th, median or 50th, 75th, and 90th? Is there something magical about these specific percentiles in statistics?*

MW: Reporting descriptive statistics by quartiles is a general way to show the range of a data set. Within MGMA DataDive, users can further customize the percentiles they want to see anything from the 10th to the 90th percentile, including the counts, mean, and the standard deviation. Each of those percentiles for a give metric and specialty are simply identifying that value's relative position in the data set, compared to the rest of the group of data.

TS: *What inferences, if any, can be made about physicians who benchmark above or below any of these commonly reported percentiles? Does benchmarking really indicate anything specific about a data point besides where it falls among all the other data points in the data set?*

MW: If physician data compared to survey data shows that they fall above or below a percentile, it only means that they rank with that population. For example, a pediatrician with compensation that falls below the 10th percentile means that more than 90% of the pediatric compensation values are higher than his or her compensation. You would have to dig deeper into reasons why he or she is earning low compensation, beyond just looking at the survey data. The benchmarks are just the starting point.

TS: *Many industry participants seem to be suspicious of a physician making in excess of the 90th percentile of MGMA or other survey data. Is there anything suspect, anomalous, or indicative of overpayment simply because a physician benchmarks about a given percentile, including the 90th?*

MW: We have an "out-of-range" editing condition that is triggered when folks report that this physician is earning a high compensation, say more than $1 million. We reach out to that participant and say, "Did you accidentally enter in an extra zero here? It looks like this particular physician is making a lot of money."

Actually, in some cases, extremely high salary is not a mistake. They go on to explain the different responsibilities and roles that the particular physician has to fill that necessitates them to be earning compensation that is much higher than their peers. While it is rare, that is just inherent in being part of the 90th percentile. They are a very small subset of the population, but it is not unheard of, and it does happen. This emphasizes the point that there can't be a blanket statement of "this is what is necessary for everyone," across the board. Appropriate metrics and percentiles are always going to be based on the situation for the particular practice and the particular physician.

4.0 The Relationship Between Production and Compensation in the Data

TS: *For several years now, the marketplace has been setting compensation based on production levels, primarily using wRVUs, but also professional collections. From MGMA's perspective, what relationship exists between compensation and production?*

MW: Generally speaking, the more providers earn in wRVUs and collections, the more they earn in compensation. However, this isn't always the case. I alluded to the fact that someone might be earning more compensation but be producing fewer work RVUs or collecting less than their peers. There could be items such as directorship responsibilities or oversight on different administrative tasks that take up their time that would factor into their compensation more than just what they are doing in a clinical capacity, just as one example. There is no one sole metric that you can utilize to say that it is the perfect predictor of compensation. It is always going to be a mix of things.

TS: *A commonly used method for establishing physician compensation, particularly for FMV purposes, is the so-called percentile matching method. Under this method, a physician is paid a level of total compensation commensurate with his or her benchmarked level of production using survey data. For example, if a physician's wRVUs benchmark at the 65th percentile, the physician should be paid the 65th percentile total compensation level from the survey. What is your reaction to this method? Do the data have this direct level of correspondence between production, such as wRVUs, and total compensation?*

MW: It is hard to say that there is a one-to-one correlation between the two because the individual at the 65th percentile of compensation is not necessarily the same physician who is at the 65th percentile of work RVUs or collections. While you can use it as something that is informing you when you are building a compensation plan, it's not appropriate to make that one-to-one correlation and say this is exactly the same physician who is always earning or producing this many RVUs and should always be compensated at this level, first and foremost because there are different samples within the two tables. Looking at the differing provider and group counts for each specialty across the tables is a clear indicator of that.

As mentioned previously, productivity is not the only thing all physicians are compensated on, so that is another reason to not assume percentile matching is a one-to-one correlation.

TS: *A related or variation of percentile matching is the idea that, if a physician benchmarks at a given percentile for wRVUs, he or she should be paid the corresponding compensation-per-wRVU rate. In other words, if a physician produces wRVUs at the 75th percentile, he or she should be paid at the 75th percentile compensation-per-wRVU rate. Is this method correct?*

MW: MGMA data suggest that, as a physician produces more wRVUs, the physician earns more in total compensation. Conversely, as the physician produces more and earns more, *the amount of compensation per unit of wRVU actually declines.* Exhibit 1 is a snapshot of what that total compensation and total compensation-per-wRVU ratio look like across wRVU quartiles, as reported in MGMA's 2015 provider compensation report based on 2014 data.

Exhibit 1. Total Compensation and Compensation-to-wRVUs Ratio by Quartile
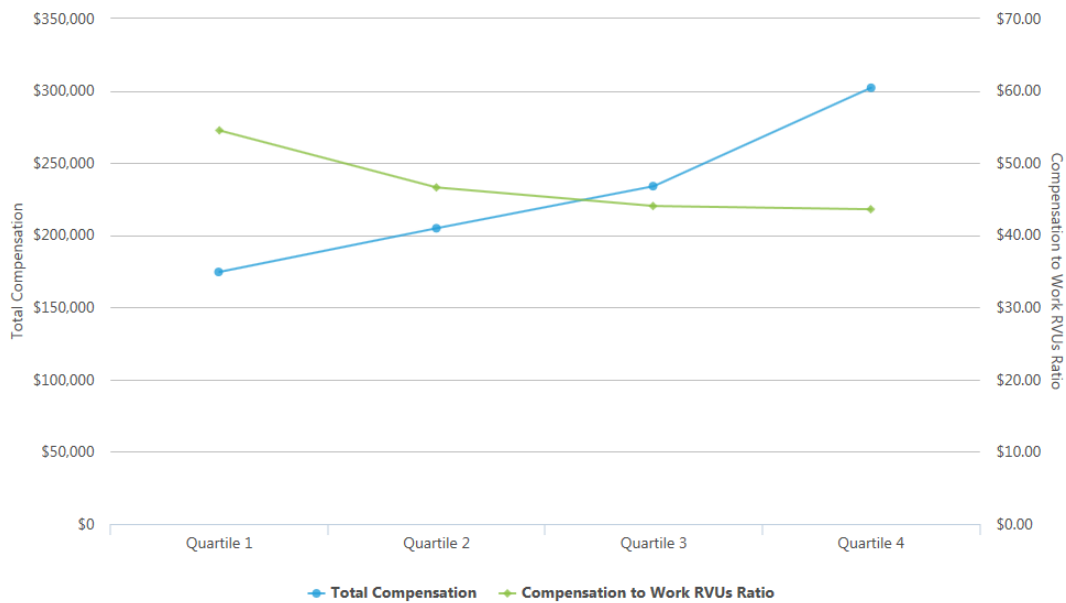
Source: Primary care specialty in Enterprise version of MGMA DataDive Provider Compensation 2015.

In this exhibit, Quartile 1 consists of wRVUs that fall between the lowest reported metric and 3,900, Quartile 2 between 3,900 and 4,800, Quartile 3 between 4,800 and 5,800, and Quartile 4 between 5,800 and highest reported metric. The physicians who fall in Quartile 1 report the lowest compensation, and, as productivity increases across quartiles, you can see total compensation increases as well.

What you'll also notice is that the ratio of total compensation to wRVU is higher for those in Quartile 1, and, as productivity increases across the quartiles, that ratio actually decreases.

Physicians with high ratios at low quartiles of productivity might be newer in a practice and are working off of some form of guaranteed compensation while they build up their patient base. And, conversely, physicians with higher productivity are likely on a volume-driven compensation plan—hence, their higher compensation earned.

We always recommend starting at the median total compensation-per-wRVU ratio figure and likely adjust down from there. You have to keep in mind that practices can't continue to pay out physicians at an exponential rate of compensation per production. If someone in the fourth quartile of productivity were paid at the same rate per work RVU as the first quartile, he or she would be making a lot more compensation, but the ability of the practice to actually pay out at that amount would be greatly reduced.

Anecdotally, I've seen that groups tend to use a tiered approach so that the ratio actually drops down lower as providers increase in productivity. Their compensation is still going up, but it is just not going to be at an exponential, and unsustainable, rate.

TS: *To further analyze the question of compensation and production, let's look at the example you developed from MGMA's Pay-to-Production Plotter tool. Exhibit 2 shows the actual data for respondents reporting both total compensation and wRVUs for the specialty of noninvasive cardiology. Why don't you step us through this exhibit and what it shows?*

Exhibit 2. Data for Respondents Reporting Both Total Compensation and wRVUs for Noninvasive Cardiology
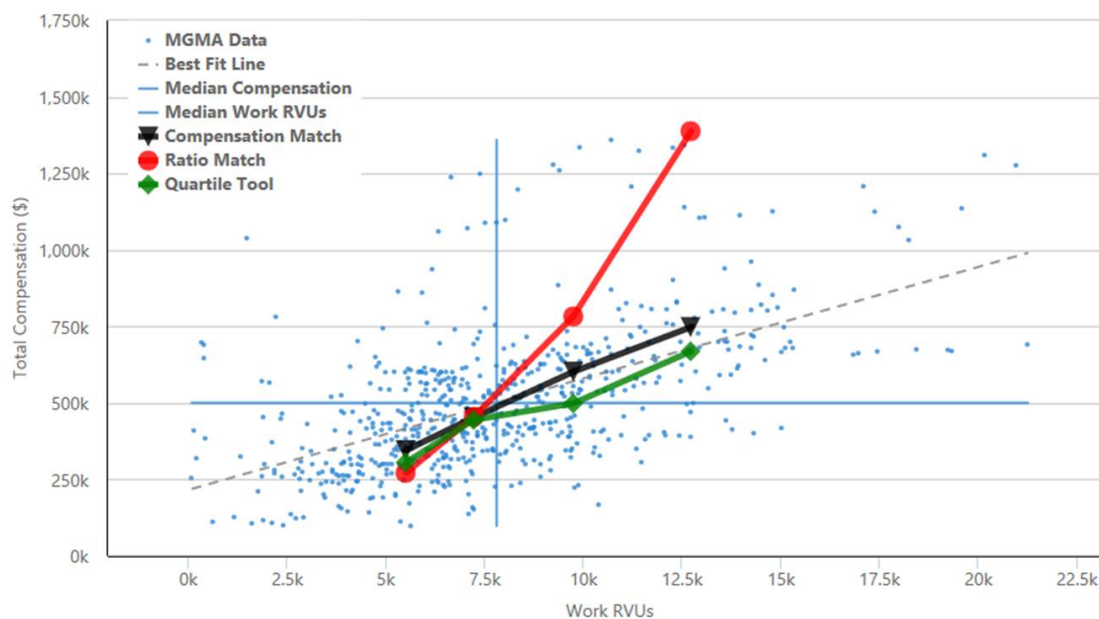
Source: Enterprise version of MGMA DataDive Provider Compensation 2015.

MW: First, you'll note the small dots scattered across the graph. Those are individual physicians who have provided both total compensation, which is laid out on the y-axis, and work RVUs, which are laid out on the x-axis. The bold horizontal and vertical lines on the graph represent the median for total compensation and work RVUs, respectively. The dashed line represents the best fit line, which is calculated from both metrics from these physicians to show their relationship. Users should look to this line as a guide and illustration of the relationship between the compensation and productivity.

The black triangle series represents percentile matching from the work RVUs table to the compensation table, based on the 10th percentiles plus four quartiles. So the first marker in the triangle series shows the 10th percentile in work RVUs plotted against the 10th percentile in compensation, and so on. You'll see that if someone were to follow this method, they'd tend to underpay and overpay a physician for his or her work in comparison to the best fit line. I'll note that the best fit line and this line matching wRVUs and total compensation look very close in this chart, for this specialty. However, keep in mind that the axis scale is in increments of $250,000, so even slight shifts can result in great discrepancies. In essence, I wouldn't recommend using the compensation matching method.

The red circle series represents percentile matching from work RVUs to the compensation/work RVUs ratio table. While the red circle series underpays physicians at first, it rapidly overpays physicians as their productivity increases. This illustrates that paying physicians a ratio rate equal to their productivity rate is not sustainable.

Finally, the green diamond series represents data from our Quartile Tool. This tool provides the best representation of physicians across productivity levels because it accounts for all the providers in a particular quartile and then shows the range of data across each of those four groups. The exhibit here plots the median compensation and work RVUs from each of the quartiles.

TS: *Let's look at these lines one at a time. Let's begin with the line matching the compensation rate with the wRVU production level based on percentile matching, which is the red circle line in the chart. Would you explain how that line was calculated and what it represents?*

MW: Often, people assume that a provider earning a certain percentile in work RVUs needs to earn the commensurate percentile in compensation-per-work RVU ratio. As I mentioned before, that's actually not what the data reflect. Generally speaking, the more a physician produces, yes, the more he or she earns in salary, but the less he or she earns per unit of RVU. But, if a user applies the matching methodology, and pays a physician who produces at the 80th percentile of work RVUs at the 80th percentile of compensation-per-work RVU ratio, the user will end up paying that physician an unsustainably high salary.

TS: *Would you step us through your analysis and opinion of this compensation model when comparing it to the actual data in Exhibit 2?*

MW: Percentile matching on the compensation per work RVU ratio underpays and quickly overpays physicians, compared to the data set. You can use the best fit line or even the green diamond quartile tool line as your guide for what most of the data are doing. There is a general pattern to the placement of the majority of data points, and that matching methodology contradicts the reported data.

TS: *Now's let discuss the line matching total compensation with production based on percentile matching, which is the black triangle line in the exhibit. Would you explain how that line was calculated and what it represents?*

MW: The compensation matching line takes physician work RVU data at the 10th, 25th, 50th, 75th, and 90th percentiles and matches them to corresponding compensation data at the same percentiles. While this method resulted in a series that more closely resembled the best fit line and quartile tool data, it is still not sound methodology. I mentioned previously that the 75th percentile of one table does not automatically reflect the same physicians at the 75th percentile in another table—the samples are at least slightly different. We always recommend utilizing something like the quartile tool, where you can look at salary and productivity data for the same physicians in a set range. That way, there is no guessing or hoping that the samples across tables are similar enough. Again, I'll note that this compensation matching line looks close to the best fit line in the sample chart due to scaling, but the reality of the data can show greater shifts that you can't predict using that methodology.

TS: *Well, let's discuss the "best fit" line. That line is based on linear regression. Would you explain the "best fit" line and what it means?*

MW: We calculate the best fit line for each specialty across productivity and compensation in the Pay-to-Production Plotter. What that line shows is the relationship between those two metrics in that given sample. It provides a general guide of how the data "behaves" across the spectrum of values. As MGMA data are descriptive statistics, we always recommend this line as a simple guide and to be used with care since our data are not from a random sample. Also, applicable filters should be put in place to most clearly understand the relationship between salary and productivity (such as organization ownership and compensation plan.)

TS: *Let's talk about the green diamond line that is based on the quartile data. Would you step us through how that line is calculated?*

MW: Quartiles are calculated for work RVUs by dividing them into four equal sections. For those physicians falling in each quartile of productivity, their corresponding compensations are calculated, and the median

compensation of each quartile is reported, along with the median work RVU in each quartile. This way we are looking at the midpoint of each of these four sections.

TS: *You've provided us with another exhibit that shows the range of compensation-per-wRVU rates for each quartile of data. I think this exhibit can help to understand compensation rates by quartile of production. Would you first explain Exhibit 3 in terms of how the numbers are calculated and how to understand the visually elements of the exhibit?*

Exhibit 3. Cardiology: Noninvasive Grouped by Work RVU Quartiles
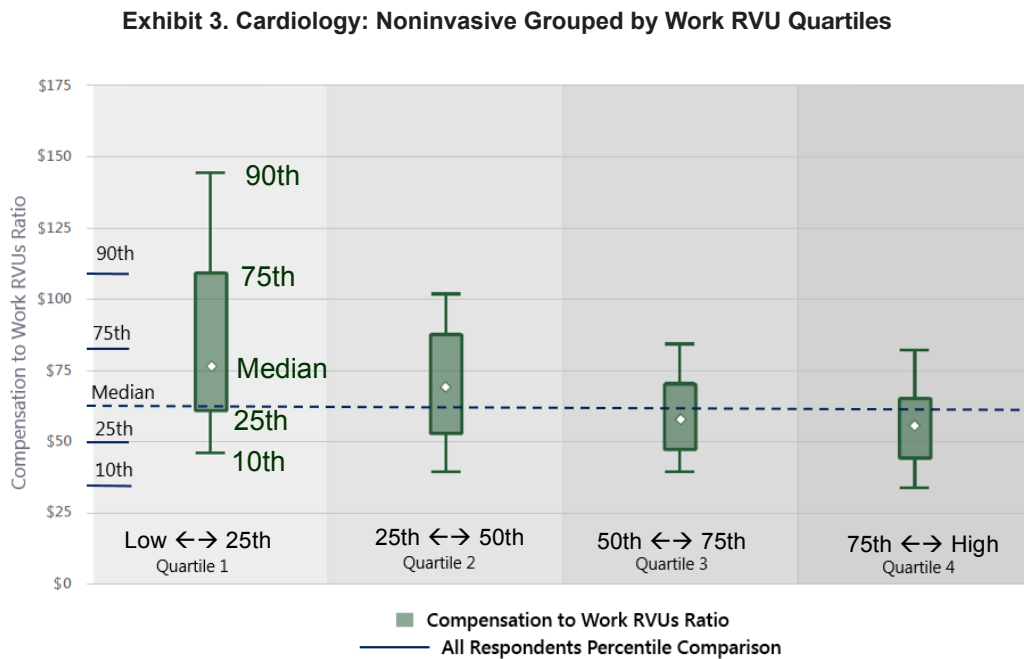
MW: Exhibit 3 illustrates the quartile tool. Here, we've grouped noninvasive cardiologists into four equal-sized groups (or quartiles) based on their work RVUs. Quartile 1 represents those cardiologists up to the 25th percentile of work RVUs (from the lowest data point up to 5,502.) Quartile 2 represents those cardiologists with 25th to median work RVUs, Quartile 3 from median to 75th, and Quartile 4 from 75th to the highest work RVU data point. Then, we've calculated the compensation per work RVU descriptive statistics for each of these four groups.

The range of compensation-per-wRVU ratios for each quartile is represented by a box-and whisker-chart. The lowest portion of the box represents the 25th percentile of compensation-per-wRVU ratio, the dot in the middle represents the median, and the top end of the box represents the 75th percentile. The whiskers below and above extend to the 10th and 90th data points, respectively.

The lines on the y-axis represent the 10th through 90th percentiles for all respondents, regardless of quartile group. Using this exhibit as an example, you can see that those producing under the 25th percentile in work RVUs in Quartile 1 are actually reporting higher compensation per work RVUs than those who produce more work RVUs. Also, the 90th percentile of compensation-per-work RVU data for all respondents is only reflected in those physicians performing the least amount of work. This tends to debunk the percentile matching theory of data across tables when it comes to productivity and ratios.

TS: *The most obvious feature of this exhibit is that rates for the first quartile are really out of synch with the rest of the quartiles. The quartile median is well above the overall median for all respondents, the 75th percentile quartile rate is about*

the same as the 90th for all respondents, and the 90th percentile rate is much, much higher than the overall 90th percentile. Why is the first quartile so different?

MW: The first quartile reinforces the caution that should be taken when using anything beyond the median of compensation-per-work RVU ratio data. Physicians who earn a normal salary but produce less in work RVUs are often subsidized by their practice perhaps because they are building out their panel size. As such, the resulting ratio of compensation to work RVU is going to be high, just due to the math.

TS: *We've discussed two exhibits that look at the specialty of noninvasive cardiology. Would you expect for the data to be similar with other specialties? What other data patterns might we expect to see?*

MW: Data need to be analyzed on a case-by-case basis. I would expect the general behavior of data to be the same across like specialties, but each specialty type has its own nuance in how much productivity is produced and how those physicians are typically paid.

5.0 Common Misuses of MGMA Survey Data

TS: *Are there any misuses of MGMA data that you commonly observe in working with users of the data? What are the top misuses and why are these uses not appropriate?*

5.1 Dividing Across Tables/Calculating Ratios Using Different Tables

MW: A common misuse of our data specific to the *Physician Compensation and Production Survey* or the *Provider Compensation* survey in DataDive is dividing across tables to achieve ratios. Our tables have different populations. You are going to see, for example, that the physician compensation tables are always going to have larger counts than any of the productivity tables because physician compensation is a required question. We are always going to have more physicians in this table.

You can look at the comparison. Collections, for example, the differences in the provider and medical practice count—there are two different samples and two different populations within these tables. If you are dividing descriptive specifics across these two samples, you are not comparing like information. We actually create MGMA ratio tables, so we will have tables that already say compensation per collections or compensation per work RVU.

We are calculating that on the individual basis for each individual provider and then compiling the descriptive statistics. We always urge folks to utilize that to inform them instead of developing their own ratio table.

5.2 Misuse of Data From the Cost Survey

MW: Another common misuse is how data from the *Cost Survey* or *Cost and Revenue* module from DataDive is used for benchmarking. We always urge folks to use that when thinking about the practice as a whole, but, if they are benchmarking individual physicians, don't use the per-FTE data—go back to the physician compensation report because it is definitely going to be the best resource for individual.

A common misuse of the *Cost Survey* is that folks tend to add up the subcategories in the staffing tables. You will see in the exhibit on the bottom right that folks might try to add up those different ancillary support staff and wonder why it adds up to 0.91 instead of the 0.73 you see. For one thing, we don't always get all of our response rates for all of those different support staff types, and there might be outliers there as well that need to be trimmed out. If folks need to utilize total support staff lines, we always suggest they use the total line and try not to add between the different categories.

Something else that happens is that folks will divide across tables in the *Cost Survey* and the *Physician Compensation and Production Survey* or the online version, the *Provider Compensation* survey in DataDive. Those are definitely two separate survey populations. They can be informative or interesting to look at, the different benchmarks from each survey, but you can't derive direct correlation from collections in the *Cost Survey* to collections in the *Provider Compensation* survey data. The *Cost Survey* is going to include the technical component in collections, as well as monies related to any drug charges, whereas, in the *Provider Compensation* data, collections will only reflect the providers' professional contribution.

5.3 Use of Charges

MW: Another common misuse of MGMA data is relying on professional gross charges. We do still include those tables within the book and DataDive. They are interesting to take a look at and maybe stay informed, but they are always going to vary between practices. We always recommend that folks utilize something that is more consistent and standardized such as RVUs or collections as physician performance benchmarks because what one practice is charging is never going to be the exact same as another practice down the street.

There are definitely more consistent benchmarks found within the relative value units or professional collections.

TS: *I often see valuators and other market participants claiming operational or other problems with physicians who benchmark at the 25th percentile for this or that metric. Their only evidence of such problems is benchmarking levels. For example, I once saw a valuation report that concluded a practice has revenue cycle problems simply because collections per wRVU benchmarked around the 25th percentile. I heard another valuator say he has never met a 25th percentile physician. What do you say to those who think low benchmarking alone is proof of operational or other problems with a physician or practice?*

MW: In your example, it's true that the particular practice does not have as much revenue as others within our survey sample, but only looking at one benchmark hardly tells the entire story. Anyone using benchmarking data has to look at multiple facets of the situation. Is the level of expense per wRVU also low? If so, there might not be an issue at all. Also, what does the payer mix look like? What do other practices in that market look like? As we've mentioned previously, benchmarks are pieces in a larger story. It is necessary that survey users take into account influencing characteristics specific to each provider and practice they are evaluating.

6.0 MGMA's Quality Review Processes in Publishing Data

TS: *MGMA follows specific procedures every year in reviewing data submissions to ensure the integrity and quality of the data. Would you step us through your processes?*

MW: We begin the data editing process as soon as a participant enters information into the survey, in real time. We give them the option to review their information before submitting and let them know if we identify missing required values, data that defy logic (e.g., work RVUs that are greater than total RVUs), and data that are out of expected range (e.g. physician earning $2 million). Once the participant submits the survey, any outstanding data edits are run and sent to them for their corrections. We follow up via email and phone to clarify responses, if we don't hear back. After we've received all of the responses, we'll run a statistical formula over the data set to automatically remove high and/or low outliers. Next, we manually walk through the data set to identify any further data that skews the data up or down. We take six to seven levels of data analysis to ensure the accuracy of the information we're reporting.

TS: *Earlier in the interview, you mentioned data trimming. Would you explain what you mean by data trimming and what are the criteria for trimming data out of the data set?*
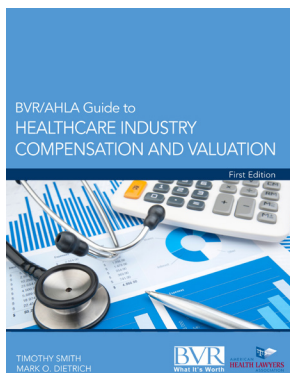
MW: Again, we apply a statistical formula that trims high and/or low outliers from a data set. After that, upon manual review of the data, if we view data at one extreme or the other, or even throughout, that does uphold the integrity of the data set, we can opt to suppress that information. We look to shifts in medians from year to year and the standard deviation relative to the data set to help guide our analysis.

TS: *Over the years, I've heard some valuators criticize the metric collections from professional charges as being inaccurate because it really includes total collections, including those from technical component services. These valuators say practices are reporting total collections rather than professional collections. The reason they make this claim is that the physicians they value all have collections that benchmark low in comparison to wRVUs. Also, many practices state they cannot calculate professional collections. Does MGMA run any quality checks on the data reported for professional collections?*

MW: MGMA runs a variety of validations on survey information submitted. We request only professional collections and total RVUs with no technical component included, as the compensation survey's goal of the *Provider Compensation* data is to measure a provider's professional performance. We also require a designation of level of technical component included in productivity, as a double check on the data. As a result, the standard MGMA collections and total RVU tables only include data with 0% technical component included.

We also have validations on the range of numbers, so, if someone is selecting 0%, but they are providing very high collections or total RVUs, we have a data edit that asks participants to clarify the figure.

TS: *Meghan, thank you very much for taking time to address these questions. We very much appreciate you providing insight and feedback on usage of MGMA's data.*