RESPONSE TO BSG Analytics WHITE PAPER

August 17, 2020

BHCG Response to Facilitate Physician and Health System Performance Improvement Efforts:

The BSG Analytics White Paper Regarding the BHCG/GNS Wisconsin Physician Value Study

I. EXECUTIVE OVERVIEW

The Business Health Care Group of Wisconsin (BHCG) commissioned the Physician Value Study in 2019 to support its goals of increased value and continuous improvement of health care in Wisconsin – with a broader goal articulated by member company CEOs of making Wisconsin a great place to do business. In short, BHCG believes that excellent care can be available in the state at a more reasonable cost by reducing (or eliminating) activities that do not provide good health care value, while paying properly for those health care actions that do provide appropriate, high quality care. Wisconsin health care has consistently been shown to be both high quality and high cost. There is in Wisconsin a clear opportunity to reduce costs while maintaining and even improving quality. BHCG is committed to this outcome.

The study was commissioned with a clear understanding of the requirements for a study focused on the goals articulated above. This included the ability to address two key issues:

1. **Data and analytical tools**: Assessment of the value provided by specific health care providers in Wisconsin has often been impaired by a lack of data availability and analytical tools. The BHCG study addresses this issue by using an industry-leading analytics firm, GNS Healthcare (GNS), utilizing the most advanced artificial intelligence techniques, paired with the broadest database available in Wisconsin – the database of the Wisconsin Health Information Organization (WHIO), the state's all payer claims database. Further, the study was led by experienced practitioners with many years of experience in analysis of health care quality and cost effectiveness in multiple geographic regions.
2. **Actionable and meaningful outcomes**: The outputs of studies like BHCG's must include clearly described and feasible solutions with meaningful real world outcomes rather than just abstract prescriptions. The BHCG study, consistent with many other studies, demonstrated a lack of correlation between the quality of care provided and the cost of achieving the outcomes of that care. The results showed that if primary care in Wisconsin was provided in a manner consistent with how the top 50% of primary care

physicians (PCPs) who were most cost effective practice today, costs of care episodes led by PCPs could be reduced by over 30%. So, whether more care is steered to those higher-performing PCPs or more PCPs evolve their practices to meet those higher-performing approaches, real value can be achieved within the Wisconsin health care marketplace today without a need for radical overhaul of the system.

BHCG Physician Value Study: Potential Annual Cost Savings

	PCPs	**4 Specialist Procedures**
Total cost in study group	$1.37B	$687M
Savings by Improving Performance above 50th percentile or Steering Patients to providers above 50th percentile	**$394.5M**	**$100M**

The BHCG study design was presented to Wisconsin provider organization representatives and other content experts for feedback through in-person meetings prior to completing the analysis. The final study design and results of the BHCG study were then widely disseminated in a highly transparent manner, with BHCG being very open to additional input and commentary. BHCG's expectation was that this input and commentary would be aligned with the goals and expectations articulated above, namely, increased value and continuous improvement of health care in Wisconsin and making the state a great place in which to do business. In response to our study, the Wisconsin Hospital Association (WHA) commissioned BSG Analytics (BSGA) to provide an assessment of the study, distributing the results of that review to many health care providers in the state.

We welcome efforts to provide constructive input and discussion, but believe it fails to meet the goals articulated above due to numerous inaccuracies and misleading statements and an effort to set an unrealistic standard for evaluations of physician and health system performance, that ultimately serves to undermine, rather than advance performance improvement efforts. For that reason, we believe the BSGA report is best addressed with thorough and factual responses, which is the purpose of this document.

The BSGA report states the following:

1. *"Initiatives focused on improving health care quality and efficiency are to be commended."*
2. *"WHIO's data helps to provide direction and magnitude for potential efficiency improvements at the health system or provider group levels."*
3. *"As health care continues its transition to value-based payments, collaboration based on credible analyses and clinical insights, will foster continued improvement and properly align incentives to ensure the focus remains on quality, patient outcomes and efficient health care delivery."*

BHCG heartily agrees. That is why we find the numerous inaccurate, misleading and sometimes internally inconsistent characterizations of the BHCG Physician Value Study troubling. Here is a sample of such characterizations and our comments regarding them:

1. While noting that WHIO data can be a good source for efficiency improvements, the BSGA white paper then goes on to say that it is not possible to evaluate either quality or efficiency of care using WHIO data.

2. BHCG report and national organizations that endorse measures for evaluation of quality of care, as well as experienced analysts of provider performance and most provider organizations, believe evaluation of individual physician performance provides highly actionable information that is probably more useful than anything else in efforts to improve quality of care. In contrast, the BSGA report asserts that assessment of individual physician performance should not be the basis for health care performance analysis and improvement. Further, and without any supporting rationale, it states that attempts to evaluate quality of care delivered by primary care physicians (PCPs) *"may promote the fragmentation of care, which can lead to the duplication of tests, inconsistent care and a lack of coordination that would adversely impact outcomes."* This assertion is the opposite of the likely outcome of properly structured analyses such as the Physician Value Study, which reviews episodes of care managed by PCPs and highlights inefficiencies due to such things as duplicate services and lack of well-coordinated care.

3. The BHCG study uses statistically valid and contemporary data from a very large data repository (WHIO), consistent with the approach used by many national quality review organizations. BSGA's review asserts that valid evaluations of quality of care require at least 3-5 years of observation and therefore a one-year time frame is too short for credible analysis.

   In fact, most analyses of quality of care in the U.S. employ a one-year timeframe. Use of 3-5 years of observation actually has the effect of "smoothing" or washing out opportunities for improvement that have occurred during the study timeframe. The ideal approach is to repeat the analysis at defined intervals to validate the replicability of the results and trend the results over time. BHCG intends to do exactly that, using the results of this study as the baseline as we prepare for the next iteration of this study. Continuing delay associated with gathering and incorporating long periods of observational data would undermine the ability to make results actionable in any immediate timeframe.

   In multiple meetings held by BHCG with health systems, industry groups and employer representatives, study methodology and results were described in detail. The BSGA

report, however, refers to the BHCG study as using a "black box" methodology, without noting what additional study descriptions would make the methodology more transparent. Nevertheless, BHCG is happy to provide additional study detail, as appropriate, to address this type of concern.

4. The BSGA report asserts that, for an evaluation of physician (or health system) performance to be valid, it must examine the "totality of medical care," which it defines as:

- Care provided by all types of physicians, as well as physician assistants, nurse practitioners and physical therapists and other professionals; AND
- in addition to evaluating physicians on the basis of widely accepted evidence-based measures of quality of care, it also must assess physicians' clinical insights, interpretation of lab results and treatment plans, patient health status over time; AND
- the efforts of integrated delivery systems.

We are unaware of any study that satisfies these requirements and do not believe it is even feasible to satisfy them or to make the results of such a complex study actionable. Moreover, the National Committee for Quality Assurance and the Centers for Medicare and Medicaid Services both use the measures of quality of care we employed in our study. We are concerned that BSGA's proposed "requirements for valid provider performance evaluation" are deliberately set at a level than cannot be achieved in the real world and would simply leave the inefficiencies of the status quo in place, maintaining the high cost system currently prevailing in Wisconsin.

Three things are required to improve health system performance: a) information regarding health provider performance at a level that is actionable; b) a strong commitment on the part of health care providers to improve their performance; and c) payer commitment to reward higher performers and to rebalance payments such that providers do not lose excessive revenue as they transition to value-based payment. The BHCG study produced actionable information. The BHCG member employers are committed to paying for value.

The remainder of this document provides our detailed response to the WHA/BSGA assessment.

II. PURPOSE AND METHODOLOGY

BHCG's top priority is to assist its member employers to purchase high value health care. To help achieve this objective, BHCG undertook a study in 2019 to evaluate the value (quality and cost) of care provided by certain types of physicians in Wisconsin. The study was funded by BHCG and the Greater Milwaukee Business Foundation on Health with the hope of 1)

supporting physician and health system performance improvement and 2) informing employer benefit design decisions in an effort to purchase higher value health care.

After a thorough search, BHCG selected GNS Healthcare (GNS), a leading provider of artificial intelligence applications in health care, to conduct the study using claims data from the Wisconsin Health Information Organization (WHIO) – the largest source of health care information in the state, including data on four and a half million covered lives and $100B in charges.

The objectives of the study performed by GNS were to answer the following questions:

1. What is the quality and efficiency of each individual primary care physician (PCP)?
2. How does quality and efficiency vary across "practice groups?"
3. What is the savings potential of moving patients to higher efficiency PCPs and/or improving the performance of lower performing PCPs?
4. What clinical care patterns differentiate higher and lower performing PCPs?
5. What is the savings potential in select specialty care areas?

The study analyzed WHIO claims data from 2016-2017 and employed evidence-based medicine (EBM) measures of quality of care that are routinely used by the Centers for Medicaid and Medicare Services (CMS), the National Commission for Quality Assurance (NCQA) and/or The Joint Commission (TJC).

In performing the study, GNS went to great lengths to overcome a shortcoming of many analyses of the quality and cost of care delivered by individual physicians that is particularly, and justifiably, bothersome to health care providers, i.e., that the number of observations used to estimate the quality and/or efficiency of individual physicians often is not high enough to enable estimates to be made with "precision." That is, the confidence interval around an estimate of the quality or efficiency of care delivered by an individual physician is often broader than what physicians consider appropriate. Because of that typical shortcoming, GNS limited its analysis to only those PCPs for whom there were >=100 observations. Many analyses of physician performance require only >=30 observations. In order to achieve the increased precision that physicians desire in an evaluation of their performance, it excluded PCPs who did not meet the requirement for the higher number of observations.

Specifically, GNS evaluated the performance of 3,760 PCPs out of the 4,009 PCPs for whom there was complete data in the WHIO database. A total of 666,636 completed episodes of care provided to 456,753 patients were included in the analysis of the 3,760 PCPs. As importantly, the BHCG study also reported the results of a sensitivity analysis that included the number of PCPs whose performance could be evaluated at different levels of confidence – something that is rarely done in analyses of physician performance.

GNS focused its analysis of a PCP's performance on the care provided for the 26 diseases, and types of preventive care requirements, that PCPs deal with most frequently. The most commonly used evidence-based measures of quality of care for these conditions specify in detail the characteristics of patients to whom each measure applies. As a result, there is no need to – and other organizations do not – risk adjust performance on these measures to account for variation in patient demographics, severity of disease, or comorbidities. When estimating "efficiency of care," however, one does need to adjust for factors outside of the physician's control that could influence efficiency of care. Efficiency of care was analyzed using "normalized" prices that are adjusted for geographic differences in prices because WHIO is prohibited under state law from disclosing the actual amount paid to a particular provider by an insurer for a service. Given this constraint, GNS analyzed the "normalized" amounts paid, which reflect an average amount paid for a particular service.

In its analysis, GNS used a proprietary artificial intelligence (AI) platform to adjust for age, gender, specific diagnoses, number of diagnoses, disease severity, complications and line of business (i.e., Medicare, Medicaid and commercial insurance). GNS compared the individual patient ratings of disease severity and complications that the OptumInsights' software creates (and is embedded in the WHIO data) and found they did not improve the accuracy/performance of the efficiency model that GNS produced using its own AI platform.

Finally, GNS estimated the amount of money that could potentially be saved if physicians whose case mix-adjusted efficiency was in the lowest 50% of all the physicians' case mix-adjusted efficiency practiced like the physicians whose case mix-adjusted efficiency was in the highest 50% of the case mix-adjusted efficiency distribution. Potential cost savings estimates were produced not only for PCPs, but also for interventional cardiologists who perform PTCA, orthopedic surgeons who perform hip replacement, orthopedic surgeons who perform knee replacements and obstetricians' performance with regard to deliveries.

The total cost of care provided for the 26 conditions examined in the PCP analysis was $1.37 billion (in 2017 dollars). GNS estimated that $394.5 million (35%) of that total could potentially be saved if PCPs whose estimated efficiency was in the lowest 50% of scores practiced like the PCPs in the highest 50% of scores. This estimate is consistent with results of many other analyses that have estimated that about one-third of the cost of health care in the United States is unnecessary. GNS also estimated that $100 million (15%) could potentially be saved if specialists included in the study who performed in the lowest 50% of estimated efficiency scores performed like their peers in the top 50% of scores.

Three things are required to improve health system performance: a) information regarding health system performance at a level that is actionable (i.e., that identifies a specific component of care provided to a specific patient that is inconsistent with widely accepted best practice), b) a strong commitment on the part of health care providers to improve their performance, and c) payer commitment to reward higher performers and to re-balance

payments such that providers do not lose excessive revenue as they transition to value-based payment. The BHCG study produced actionable information regarding how to improve the quality of care delivered for the most common conditions cared for by PCPs and BHCG provided health systems with measure-specific scores for each of their physicians that were included in the study.

Following the public release of the BHCG study, BSG Analytics, LLC (BSGA) issued a white paper entitled *"The BHCG/GNS Wisconsin Physician Value Study: Structural Issues Undermine its Effectiveness for Physician Referrals and Health System Assessments."* There are several statements in BSGA's white paper with which we agree. However, BSGA's white paper also contains numerous assertions that are inaccurate and proposes "requirements" that no analysis of quality and cost of care can realistically fulfill today or in the near future. In an effort to address these shortcomings in BSGA's white paper, we offer the following responses.

III. ANALYSIS

A. Assertions With Which We Agree

1. *"Initiatives focused on improving health care quality and efficiency are to be commended."*

   We are in complete agreement.

2. *"The [BHCG] study's quality measures focus on process measures, rather than outcomes. The OptumInsights quality measures used in the analysis are only based on information that is included in claim forms. While the OptumInsights quality measures may be helpful in identifying whether physicians are appropriately prescribing diagnostic tests and medications, they are not a good tool for measuring improvements in a patient's health status over time."*

   We agree the quality measures used in the BHCG study are helpful in identifying whether physicians are appropriately prescribing diagnostic tests and medications. Virtually all quality measures for primary care that are used by CMS, NCQA and TJC are process measures. A few intermediate outcome measures examine how well something like HgbA1c or blood pressure is controlled. Since there are no "industry standard" quality measures for primary care that look at a patient's health status over time, the measures we employed to determine if physicians are appropriately prescribing diagnostic tests and medications are based on well-documented evidence that shows that compliance with these clinical practices reduces the likelihood of future adverse events. Reviewing a patient's health status over time, as a measure of clinical quality, is a potential future opportunity, but since there are no accepted measures, BHCG would

have needed to create them for this study. Thus, we utilized only those measures that are currently widely accepted as quality measures.

3. *"Physicians were evaluated primarily on whether they ordered a test rather than their interpretation of the results and subsequent treatment plan."*

   Similar to #2 above, we are not aware of any industry standard quality measures that look at how a physician interprets a test result or how such interpretation influenced their subsequent treatment plan. The quality measures we employed measure a physician's compliance with clinical practice guideline.

4. *"WHIO's data helps to provide direction and magnitude for potential efficiency improvements at the health system or provider group levels."*

   We agree. In fact, the WHIO data is the only source of data in Wisconsin to assess care across the continuum of care with a very large sample size and breadth of data. BHCG member employers are committed to expanding the robust WHIO database to make it even stronger as a source for future analysis.

5. *"Several of the conditions included in the study are chronic diseases that require significantly more medical resources in the first year of care to mitigate health care utilization in subsequent years."*

   We are not sure what BSGA means by "*significantly more* *medical resources,"* but we agree the chronic diseases for which we evaluated care do require resources to be expended in order to reduce future utilization. That said, the tests and treatments included in the quality measures we employed are not expensive, especially in comparison to the cost of avoidable future care, such as hospitalizations. If BSGA is suggesting certain physicians treat patients in the first year of diagnosis while other only treat patients in subsequent years, it is an unusual conjecture and does not seem consistent with real world practices.

6. *"The study uses "black box" technology that cannot be independently validated. Risk adjusting is critical to normalizing the medical utilization used to treat patients so that physicians are not penalized for having sicker patients. The risk-adjusting methodology used in most WHIO analyses relies on Optum's severity adjustments, which are well documented and transparent. GNS Healthcare's Reverse Engineering Forward Simulation (REFS) determined the relative weight of the study's risk-adjusters based on its machine-learning algorithms."*

   We agree that *"risk adjusting is critical to normalizing the medical utilization so that physicians are not penalized for having sicker patients."* That is why we employed

extremely sophisticated artificial intelligence to perform risk adjustment. We also checked to see if OptumInsights' severity adjustments improved the fit of our models beyond the adjustments we employed. They did not improve them. While it is unlikely most people have the analytical training and education to understand the mathematics and statistics employed in the type of machine learning we employed, the methodology was clearly articulated – and not a hidden, "black box" approach. The methods we employed are based on the work of Judea Pearl, who won the Turing Award (the "Nobel Prize" of computer science) for his work on causal learning. Dr. Pearl's work is described in a bit more understandable terms in *"The Book of Why: The New Science of Cause and Effect,"* co-authored by Dr. Pearl and Dana MacKenzie.

7. *"The methodology weights each physician's performance based on his or her unique mix of episodes. Asthma will have twice the weight if physicians primarily treat commercial or Medicaid patients than if they treat Medicare patients. Conversely, ischemic heart disease is twice as important for providers who treat Medicare patients as it is for those who treat commercial or Medicaid patients."*

   We believe this is the proper methodology to employ as it weights each PCP's score based on the types of patients each PCP cared for, not the diseases that a "typical" PCP is most likely to encounter based on the type of insurance a patient is covered by.

8. *"As health care continues its transition to value-based payments, collaboration based on credible analyses and clinical insights, will foster continued improvement and properly align incentives to ensure the focus remains on quality, patient outcomes and efficient health care delivery."*

   We hope the *"will"* in this statement is true and that health care providers are "willing" and committed to improving the quality and cost efficiency of health care in Wisconsin, in partnership with all payers, for the betterment of the people of Wisconsin.

B. Inaccurate Assertions With Which We Disagree

1. *"Efforts to provide meaningful and actionable information for health systems and employers … must also reflect Wisconsin's highly integrated health care systems."*

   We disagree with this assertion. As licensed practitioners, actionable information that helps an individual physician know that he/she is not complying with widely accepted evidence-based guidelines enables performance improvement. Moreover, as stated above, the way we calculated quality scores gave PCPs credit for tests and treatments ordered by other clinicians, in addition to those ordered by the PCP him(her)self. As a result, our results reflect the performance of Wisconsin's highly integrated health care systems as much as they reflect the performance of individual PCPs.

2. *"It (the BHCG study) does not address the appropriateness of care, referral patterns and coordination of care within integrated health systems."*

   The BHCG study was not intended to address these dimensions of care within a health system as an interrelated entity – the process measures we used look at whether appropriate care was provided by the physicians being studied. Thus, a high-performing physician in an integrated system that performs near the average level is not disadvantaged by system-wide measures.

3. *"By focusing on just 26 of the more than 445 Base ETGs attributed to primary care physicians, the analysis excludes two-thirds of the episodes and 60 percent of the medical resource utilization attributed to family practice physicians, internal medicine physicians and pediatricians who are included in the study."*

   We are not sure how BSGA calculated these proportions/percentages. It is true that we focused on the 26 chronic conditions that are the most commonly treated by the PCPs in the study and used the ETGs related to these conditions. This implied criticism seems like "damned if you do, damned if you don't." Had we instead looked at conditions seen infrequently by PCPs; we likely would have been criticized for doing so.

4. *"Analyses limited to individual medical specialties fail to consider the successful efforts of integrated health delivery systems to improve the totality of the medical care they provide."*

   While this is true, it is highly misleading as illustrated by this simple analogy. Evaluations of toxic substances in individual streams and lakes does not reflect the quality of all water on our planet. That does not mean that testing water in streams and lakes for toxic substances is not beneficial and should be discontinued. Similarly, evaluation of the quality of care provided by physicians in a particular specialty is beneficial. We are not aware of any study that has ever been performed that evaluates "the totality of the medical care" and doubt that could be done. (See our comments below in the Unattainable Requirements section.)

5. *"In addition, they (analyses limited to individual medical specialties) may inadvertently promote the fragmentation of care, which can lead to the duplication of tests, inconsistent care and a lack of coordination that would adversely impact outcomes."*

   This is absolutely not the case. Evidence-based measures of quality of care are based on the results of widely accepted peer reviewed studies and reflect what are considered to be best clinical practices. There is no reason to believe evaluations of the quality of care delivered by individual medical specialties would promote fragmentation of care and

erode coordination of care. Moreover, virtually all, if not all, studies of quality of care examine physicians within a single specialty (or similar specialties, in the case of primary care). If BSGA's statement were true, then previous reports of high quality care provided by Wisconsin's care delivery system, as measured by evidence-based quality measures, would also suggest the care provided by Wisconsin's care delivery system is contributing to duplication of tests, inconsistent care, poor care coordination and poor patient outcomes for years.

6. *"The recent report released by the Business Health Care Group (BHCG), the Wisconsin Physician Value Study … seeks to use the study's results to help employers steer patients to high-performing primary care physicians."*

   This statement is misleading. In fact, in the multiple presentations we made to various provider groups in Wisconsin, we indicated it is <u>not</u> practical to steer patients to only the highest performing physicians because those physicians do not have sufficient capacity to care for all the patients cared for by lower performing physicians. For that reason, the primary goal of the BHCG study is performance improvement, which would result in a larger pool of high-performing physicians, which we believe health systems could accomplish with a steadfast commitment to this goal.

7. *"Basing a health system score on the performance of primary care physicians overestimates their ability to control the medical care provided outside of their offices."*

   We partially agree with this statement, but it doesn't really apply to our study. The methodology we employed gives PCPs credit on quality scores for services provided by any provider – not just the PCPs. As a result, PCPs' quality scores benefited from care provided by other physicians, physician's assistants and nurse practitioners. The frequency with which PCPs refer patients to specialists, and the quality/efficiency of the specialists to whom PCPs refer patients, are under the control of the PCP. If the BSGA argument is PCPs are somehow required to refer to inappropriate or inefficient practitioners for "care provided outside of their offices," then that is, indeed, an issue that should be addressed to enhance the value of care received in Wisconsin.

8. *"Even within its limited focus on primary care, the study was unable to confidently determine whether three-fourths of the physicians evaluated performed above or below the state average. That ambiguity makes it ineffective for referrals or performance evaluations*."

   We disagree with the idea that our study did not provide valid performance evaluations. The fact we reported on the physicians whose performance we could/could not evaluate with adequate precision is a strength of our study which should be applauded, rather than criticized.

9. *"Its one-year timeframe is too short for credible analysis."*

   This statement is unequivocally incorrect. All of the quality measures we employed in our study require only a one-year period of observation. The evidence-based guidelines on which the performance measures we employed are based reflect multi-year studies that showed the benefit over more than one year of compliance with the guidelines on which the quality measures are based.

10. *"Most analyses include three to five years of claims data."*

    We do not profess to have performed a review of all claims data-based analyses, but we strongly doubt that a) BSGA has performed such a review and b) BSGA's assertion is true. In fact, all NCQA HEDIS and Medicare STARs measures examine a one-year time period. Moreover, because of our response to the preceding item, this statement is irrelevant.

11. *"Like satellite photos, it [i.e., WHIO's data] … cannot "see beneath the trees." Evaluating quality and providing granular analysis actionable at the individual provider level requires information that is not attainable through WHIO data."*

    This statement is unequivocally false. ALL of the results of our evaluation of PCP's performance are actionable. WHIO data is the only database in Wisconsin that has a sufficient volume of data to produce statistically valid results at the individual physician level when thoughtful business rules are applied (e.g., our requirement of a minimum of 100 observations per physician for inclusion). Wisconsin's health care systems should be performing these types of analyses themselves and providing this information to their physicians to systematically improve the value of the care they deliver.

12. *"Using incomplete data to rank physicians, as is done in the BHCG study, is unlikely to improve health care delivery but could lead to unnecessary market disruption and undermine other, more credible efforts to identify best practices."*

    We are not sure what BSGA means when it refers to "incomplete data." It may refer to the fact that we only examined care for the 26 conditions and preventive care PCPs address most frequently. We also are unsure why BSGA thinks our analysis of quality of care will "undermine other, more credible efforts to identify best practices." In fact, it was never our goal to "identify best practices." Rather, we employed widely accepted measures that reflect best practices as defined in evidence-based clinical practice guidelines to evaluate the care delivered by the PCPs included in our study. We are confident the results of our analysis can facilitate, not undermine, improvement efforts.

13. *"The study also sought to determine … the savings that could be achieved if medical care could be moved to top-tier (high-quality, low-utilization) providers as defined by BHCG."*

    As we have explained above, this statement is inaccurate. Our study estimated the savings that could be realized if patients cared for by PCPs in the bottom 50% of efficiency scores were steered instead to the PCPs in the top 50% of efficiency scores, *or*, if all of the underperforming primary care providers changed their practice patterns to perform at an above-average level.

14. *"The study also excludes nurse practitioners and physician assistants, two medical specialties that are providing an increasing share of primary medical care."*

    The performance of the PCPs whose practices we evaluated reflects services ordered for the PCPs' patients by any professional, including nurse practitioners and physician assistants, as well as other physicians. We did not intend to evaluate the independent performance of nurse practitioners and physician assistants, as they typically function as part of a PCP-led team, rather than in isolation.

15. *"Limiting health system scores to the performance of primary care physicians excludes the care provided by surgeons, oncologists, cardiologists, obstetricians, gynecologists, rheumatologists and other physician specialists; physician assistants; nurse practitioners and physical therapists. It also does not include any lab tests, diagnostic imaging or prescriptions ordered by these providers."*

    BSGA is correct, we did not attempt to evaluate the performance of all types of physician specialists. We are not aware of any evaluation of value (quality/cost) that has ever accomplished this for many reasons, including the fact that many specialties and sub-specialties do not have clinically meaningful quality metrics. The last sentence in this BSGA statement is incorrect. As BSGA itself acknowledges in several places, our study did include lab tests and prescriptions ordered.

16. *"A primary care physician who is quick to refer patients to cardiologists for hypertension may appear to be more efficient than primary care physicians who continue to treat those patients even though the physician's decision could result in less-efficient utilization of medical resources."*

    We are not sure what BSGA intended to say here, but the truth could just as easily be the opposite of what is stated here. That is, a PCP who is "quick" to refer patients to a cardiologist for hypertension could be found to be more inefficient, not more efficient. Regardless, what we measured was PCP efficiency in utilization of medical resources. Whether PCPs who are "quick referrers" are more or less efficient than "slow referrers" could be the subject of a separate study.

17. *"The BHCG analysis does not identify any of the quality measures it used to determine quality scores. This makes it impossible to validate the appropriateness of the quality measures tied to each of the 26 conditions included in the analysis or to determine the areas where physician quality could be improved."*

    BHCG made it clear it would provide a list of all the quality measures employed in its study. If BSGA wanted the list, they could have requested it. Moreover, BSGA surely is familiar with the quality measures employed by CMS, NCQA and TJC for the diseases/conditions we examined. In fact, BSGA's white paper suggests in several places it is familiar with Optum's EBM measures.

18. *"Comingling payers prohibits "apples-to-apples" comparisons."*

    We are not sure to which "apples" BSGA is referring, but we do not believe comingling data from different payers prohibits meaningful comparisons. Most PCPs care for patients insured by Medicare, commercial insurers and Medicaid. We acknowledge Medicaid patients may be subject to more social determinants of health (e.g., poor economic situation and inadequate transportation) than people covered by Medicare or commercial insurers. Moreover, excluding patients with a particular type of insurance, or trying to stratify our analysis by payer, would have resulted in sample sizes that are too small to examine accurately. Our study used WHIO data, which is sourced from multiple payers. The data is submitted in a standard format, integrated at the patient level and is evaluated for data quality at multiple steps in its processing. Each patient in the dataset has a unique and persistent ID and each physician is identified using a standardized provider registry.

19. *"The expected savings in medical utilization are likely overstated. The WHIO database uses a "standard price"*[i] *established by Optum, which provides the episode groupers. The standard price is designed to mitigate differences in negotiated discounts or the cost of living in geographic areas: $1 in standard price in Milwaukee equals $1 in standard price in Green Bay or La Crosse. This allows the standard price to be used as a proxy for utilization. The "savings" cited in the BCHG study are identified as dollars, but are, in reality, a reflection of a reduction in utilization. Reducing inappropriate utilization can impact costs, but the impact will vary by provider and geographic area depending on the actual rates charged for individual services. Reducing inappropriate utilization in a low-cost area will have less financial impact than reducing inappropriate utilization in a high-cost area."*

    We agree with most of this statement, but we believe the standardization of prices may just as easily have <u>understated</u> the potential savings, particularly for providers whose patients use facilities that a) have very large market share or b) are institutions

considered to be "must have" in-network – because such facilities are able to negotiate much higher payment rates. In fact, the standardized price likely underestimates potential savings for commercially insured patients and overestimates them for Medicaid and Medicare patients.

C. Unattainable Requirements

1. *"Integrated health systems should be evaluated on the totality of medical care provided by the full spectrum of health care professionals – primary care physicians, specialists, physician assistants, nurse practitioners, physical therapists and other professionals."*

Evaluation of "the totality of care" would be ideal, but it has never been done before and is not feasible anytime soon. We can't think of a reason that an evaluation of the totality of medical care needs to be performed in order for an evaluation of physician performance to be valid or credible. Suggesting that it does likely would preclude any provider performance evaluation from ever being performed. In fact, this assertion implies no provider organization should ever evaluate its own performance, NCQA should not evaluate the performance of health plans or providers, TJC should not evaluate the performance of provider organizations, CMS or other payers should not evaluate provider performance, and employers should forget about seeking information about provider performance. We consider such a view to be unacceptable.

2. *"WHIO data is insufficient to provide a complete evaluation of physician performance because it does not include clinical insights, outcomes information, lab results or electronic health records."*

We disagree. We believe WHIO data is quite adequate to evaluate the aspects of care BHCG evaluated. We are not aware of any study that has evaluated "clinical insights;" moreover, we are not sure how one would conduct such a study. We also believe analysis of electronic health records would not be sufficient to get a complete picture of patient outcomes or the care patients receive, as the electronic health records of a provider organization rarely includes the care provided by out-of-system providers or the cost of care. In addition, analysis of clinical insights, outcomes information, lab results or electronic health records would not examine "the totality of medical care provided by the full spectrum of health care professionals." We thus again suspect this criticism is intended to set "requirements for valid provider performance evaluation" at a level that cannot realistically be achieved.

IV. CONCLUSION

In conclusion, we agree with BSGA's statement, *"Initiatives focused on improving health care quality and efficiency are to be commended."* It was exactly with that objective in mind that

BHCG undertook the BHCG/GNS Wisconsin Physician Value Study. And it is in that spirit we hope the findings of the BHCG study will be used by individual physicians, physician groups and health systems throughout Wisconsin.

We also believe efforts to undermine such performance improvement by inaccurate and misleading statements should be avoided and hope responsible providers will not be distracted by them.