A young Mandarin-speaking female physician is struggling to understand why she is labeled as a low performer because of her patient service scores. She also is being held back on pay raises and may not make partnership in the medical group. Her colleagues who are White are moving ahead of her when evaluations come out. She is exhausted from taking care of a large Mandarin-speaking population who often drop in without an appointment and expect to be seen. She stays late calling patients because she doesn’t have any support staff who speak Mandarin to do callbacks for her. She was an all-A student in medical school, so her sense of failure as being labeled a poor physician has led to low self-esteem. Some of the older Chinese patients speak down to her, adding to her stress. Is this truly an underperforming physician, or is something critical being missed. The answer is in the data.

Over the past decade, data collection of diverse populations has been seen as a critical first step in addressing the growing gap of health care disparities. However, change has been slow, and resistance to even the collection of data still exists in many organizations. Even when data are collected, however, these often stay unused either because of a lack of commitment to true change or a lack of understanding of how to use the data for maximum effect. Yet one of the most valuable assets of any organization is its databases. Such databases are unique for that organization, allow for self-assessment, and offer the opportunity for strategic planning for the future.

The other key area is that of unconscious bias. The topic is avoided because it is a very uncomfortable subject; yet, we all have unconscious biases. There is increasing evidence of the impact of unconscious bias on quality outcomes by race/ethnicity, age and gender. Pay for performance for physicians and organizations is also not immune to the insidious nature of unconscious bias. Yet how can we identify unconscious bias when it is avoided as an issue, is perceived as subjective, and is extremely complex. Implicit bias testing has contributed a great deal to creating awareness of the presence of unconscious bias but has several drawbacks by being limited in its scope. An example is the issue of patient bias in the physician–patient interaction. How do we test for that? I believe the answer is in the field of analytics and big data.

Data are neutral assets that are a counterpoint to the subjectivity and emotional issues surrounding unconscious bias. Data require context and analysis to reveal their hidden secrets, whereas unconscious bias requires an objective tool to identify its presence in a constructive manner. Data allow for recognition of demographic groups in which the physician does well. But data point a laser beam at the underperforming groups in which the physician does not do well. A yin/yang situation comes into play in which the “heat” of unconscious bias is tempered by the “coolness” of the data. The physician is recognized for success, but opportunities for improvement are identified as well. When reviewing the data with the physician, we start with areas in which the physician does well and end with areas in which there is an opportunity for improvement. The discussion then focuses on what the
physician perceives may be the problem with an underperforming demographic group. This perception provides the context by which physicians are able to create strategies for improving their performance. The data allow physicians to think about different scenarios for interventions. It is also a very efficient use of a physicians’ time because they are able to focus on a specific group rather than trying to fix all groups in their panel. When the gap is closed for a specific group, the overall score will rise. For those physicians who do poorly with groups, the data suggest that general courses in service are called for or that another line of work may be in order.

Kaiser Permanente has been collecting service score data since 1992 and has broken these data down into some basic demographic groups by gender, age, and race/ethnicity. More than one million surveys per year are collected, and each provider receives an average of 100 responses per year. Data are also collected at the department and facility levels. When the existing database is used and the fields of epidemiology, analytics, and systems analysis are drawn from, information is presented in a manner allowing for sophisticated targeted interventions.

This program does the following:
1. Evaluates a physician’s panel by demographic group and performance for each group;
2. Trains providers and staff in how to read data and analyze these data tactically and strategically;
3. Allows for the physician to identify key demographic groups for improvement, an efficient use of time;
4. Identifies potential unconscious bias by either the physician or the patient;
5. Creates fairness between physicians who have large demographic groups who score lower, such as Chinese, and those physicians who don’t have such groups, when service scores are used for pay for performance; and
6. Tracks results over time to reinforce cultural competence.

What happened to the young physician who was described earlier? When the issue was identified, we took immediate steps to understand what was happening because it was a common theme with all of our bilingual physicians. We broke down the data, which led to our asking the right questions to do a root cause analysis. Her scores were comparable or better than the department average in different demographic groups. What brought her score down was having a very large Chinese-speaking population, who averages 15 points lower than White patients in the same department. It was the ratio, not her scores that gave the perception of her being an underperforming physician.

We did focus groups on our Chinese patient population to better understand their needs as well as those of our bilingual physicians. Performance metrics were changed to be based on demographic subsets as compared with the overall service score. Bilingual staff, as well as more bilingual physicians, were hired to decrease the workload. We started recognizing the value of our bilingual physicians, which was a critical step. The young physician made partner and is now one of the top-performing physicians in the group.
This program, known as the *Data, Diversity and Demographics Program*, started in 2003 with a single physician and now reaches more than 5,000 physicians in Northern California Kaiser Permanente. It allows for targeted interventions, has contributed to the improvement of our service scores with diverse populations, and has improved physician satisfaction at the same time. Our success is based on customizing our interventions to fit the needs of the individual physician and recognizing the value of each physician.