Establishing a Quality Measurement System for Cancer Care in Japan

Takahiro Higashī¹,*, Fumiaki Nakamura², Nobuhiro Saruki³ and Tomotaka Sobue⁴

¹Department of Public Health/Health Policy, Graduate School of Medicine, The University of Tokyo, Tokyo, ²Department of Healthcare Epidemiology, Kyoto University School of Medicine, Kyoto, ³Gunma Prefectural Cancer Center, Gunma and ⁴Division of Environmental Medicine and Population Sciences, Graduate School of Medicine, Osaka University, Osaka, Japan

*For reprints and all correspondence: Takahiro Higashi, Department of Public Health/Health Policy, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: higashi@m.u-tokyo.ac.jp

Ensuring the quality of care is a major objective of cancer control policy. The Cancer Control Act 2006 placed responsibility on the Japanese government to maintain the quality of cancer care nationwide. To function as centers providing high-quality care, designated cancer care hospitals (397 hospitals as of April 2012) were instituted nationwide. Although they meet the structural standards, such as the presence of radiation equipment and palliative care teams, it remains unclear whether the designation has led to appropriate provision of care and optimal patient outcomes. A national system to examine the processes and outcomes of cancer care is under development. In 2007 and 2008, the Japanese Association of Clinical Cancer Centers publicly disclosed the 5-year survival of their member facilities with strict data quality standards, including sufficient follow-up of patients’ vital status. The network of designated cancer care hospitals will follow this lead to provide a national outcome monitoring system. The processes of care have also been addressed by a government-funded research project. With the collaboration of clinical experts, 206 quality indicators have been developed for five major cancers in Japan (breast, colorectal, liver, lung and stomach) and palliative care. Each indicator described the target patients and standards of care for the patients, the provision of which was considered an aspect of quality. In 2012, the Cancer Registry Chapter of the Association of Prefectural Designated Cancer Care Hospitals instituted quality measurement using these indicators. These activities will soon lead to effective quality monitoring and improvement in Japan.

Key words: cancer – outcomes research – quality control – quality of health care

INTRODUCTION

Cancer Control Activities in Japan

Since 1981, cancer has been the leading cause of death in Japan (1). In 2010, about one-third of Japanese deaths were from cancer, and about half of men and women in Japan experience cancer at some time in their lives (2). Although the age-adjusted mortality is gradually decreasing because of a reduction in risk factors, advances in medical technology which enables early detection and effective treatment, the observed number of incident cancer cases is increasing as a result of rapid aging of the population (2), and poses a great burden on society. The need for high-quality cancer care is even greater.
This increased need led to the enactment of the Cancer Control Act in 2006 (3). The law emphasized three major areas in cancer control, namely the prevention and early detection of cancer, the guarantee of high-quality cancer care nationwide and the promotion of cancer research. The law was uniquely progressive in that it placed responsibility on the Japanese government to undertake concrete actions to ensure the quality of cancer care.

To ensure the quality of care nationwide, effective measurement systems need to be developed. Theoretically, the quality of medical care can be measured at the level of structure, process and outcome (4,5). Structure refers to the resources and environment that produce medical care, such as sufficient medical staff and provision of diagnostic and therapeutic equipment, which are necessary for high-quality care. Process measurement directly examines the care provided, either explicitly with pre-defined criteria for appropriateness or implicitly by means of peer-review by experts (6). Outcome is typically measured using factors associated with care such as survival (commonly a 5-year survival in cancer) and the quality of life (QOL) after treatment. Structure is easily measured and is thus extensively used for regulatory purposes. The government started designating certain hospitals nationwide as cancer care hospitals in 2003, and the criteria for the selection were almost exclusively based on structures, such as the presence of radiation equipment and having a palliative care team.

However, the structure does not guarantee the appropriateness of care processes or desired outcomes. To gain insight on the quality, we need to measure process and outcome. The work of establishing a system to measure processes and outcomes is currently under way. This review presents an overview of the activities to measure processes and outcomes to date and provides some perspectives for the future in Japan. Since outcomes are more intuitive than processes as indicators of the quality, we first present the methods and activities used for the outcome measurement and then discuss actions targeting measuring processes.

OUTCOME MEASUREMENT

Typically, the outcomes of treatment for cancer are measured by a 5-year survival (7,8). As simple as this indicator appears, there are a number of issues that need to be considered in the actual calculations. First, we need to fix the target patients. The easiest way is to target patients who had their tumor resected surgically, but this may miss a substantial proportion of patients, such as those with lung cancer, where more than half the patients are managed medically in the first-course treatment (9). If we are to measure and ensure the overall quality of cancer care provided to patients, a more conceptually appropriate method would be to include all patients treated. Since medically treated patients tend to be more severely ill than surgical patients because of comorbidities or advanced disease, their inclusion would decrease the 5-year survival (10). Whenever we interpret the survival data, we must pay great attention to the range of patients included.

A less apparent, but perhaps more influential factor, is the proportion of follow-up/censoring. To calculate the 5-year survival, we need to know whether patients are alive or dead 5 years after the start of observation (e.g. the time of diagnosis or first treatment). Usually, the survival calculation uses the Kaplan–Meier method (11), in which a patient known to be alive to a certain point, say at 3 years, is counted as a person at risk up to that point and is excluded from the denominator thereafter. The important assumption of this method relies on ‘non-informative censoring’, which is statistically defined as the condition that ‘a patient’s loss to follow-up does not provide any information on how long the patient would have survived’ (12). However, in the real world where missing patients are more likely to be dead than patients who continue to be followed, the assumption is not valid. A study that compared the survival calculated from hospital statistics found that the rate was higher than the true survival determined from government registration data (10).

The level of bias varied depending on the patient population and cancer types, but was as great as 19%. This underscores the importance of a thorough follow-up of the vital status of individual patients in the calculation of survival rates. A low follow-up rate (i.e. a large proportion of censored cases) can result in misleading overestimation of the survival.

When comparing the survival between hospitals, we must be aware that the patients’ baseline characteristics may also be different across hospitals (13). Some hospitals may treat more advanced diseases or patients with more severe comorbidities than other hospitals (9). Risk-adjustment methods using statistical models are frequently used to adjust for the variations in patient populations (14). Although discussion of the statistical techniques is beyond the scope of this review, we feel that the risk adjustment methods are not well enough established in cancer care. The adjustment works best when the statistical models reflect the real world, such as the variables included. Furthermore, it is known that indirect standardizations using the observed or expected ratios of survival rates, which are frequently used in risk adjustment, can result in misleading results especially when comparing patient populations whose case mix varies greatly (14,15).

PUBLIC DISCLOSURE OF SURVIVAL DATA

Public reporting of survival data by the Japanese Association of Clinical Cancer Centers (JACCC) is the first organized effort to measure and publicly disclose the 5-year survival rates of the major cancers in Japan (16,17). Member facilities calculated the 5-year survival and posted their data on the JACCC website in 2007 and 2008. The target patients for survival calculation were basically all patients, both surgical and medical, who received primary treatment at the facility. Most importantly, they paid great attention to the
quality of the data. They defined the data quality standards up-front (Table 1) and did not include hospital data which did not meet the standards. In particular, they considered the follow-up rate important because, as mentioned above, the suboptimal follow-up can result in the overestimation of survival rates. Without systematic rules for follow-up, hospitals may not be inclined to rigorously attempt to identify the vital status of patients because the true survival rates obtained from these efforts are lower than the observed (but wrong) survival rates calculated without efforts. The data quality standard therefore demanded that the proportion of censoring be kept at <10% of patients, otherwise the data were considered too biased, likely to be erroneously high, and were excluded from public disclosure.

Another unique point of the JACCC public disclosures was that they did not use model-based statistical risk adjustment of the survival data. Part of the reason was that the cancer registry data used contained limited information about patient characteristics except age and cancer stage. Therefore, both the stage-stratified survival as well as the overall survival was presented. The overall survival data were presented with the ratio of Stage I/Stage IV patients to concisely represent that the differences in patient population across facilities. This presentation helped to make the public aware that the survival data cannot be simply compared across facilities. These activities were extensively covered by the media and most mentioned the risk of making conclusions on the quality of care based on the survival data because of the differences in patient populations (18).

Recently, increasing numbers of hospitals have individually disclosed the survival data of patients treated in their facilities on their own websites. Table 2 shows the results of an Internet search we conducted using the Google search engine with the key words, ‘survival’ and ‘colorectal cancer’ in Japanese in June 2012 while writing this review. It shows a variation in the methods of calculating survival. Many targeted only surgical patients. Less than half of the reports presented the follow-up rate, making the quality of the data questionable. In addition, the differences in the severity of the patients’ conditions at admission were uncertain and so these data may not be comparable. Overall, these data were difficult to interpret in the current form. An organized structure for calculation of survival is necessary. One such structure is currently planned in the framework of hospital-based cancer registries operated in all designated cancer care hospitals. However, the way to achieve sufficient follow-up rates remain unresolved.

QOL MEASURES

Although the measurement of hard outcomes such as survival is clear cut, it may miss important aspects of the severity of the patients’ conditions. To capture these details in the outcomes of treatment, several QOL measures, both generic and cancer specific, have been developed to date. These measures have been frequently used in clinical trials and technology assessment, but have seldom been regarded as a measure of the quality of care. This is perhaps because the measurement involves questioning patients, and thus is subject to ambiguities in self-reporting and missing data because some patients are unwilling to respond. If some of the quality of care is associated with improvement in QOL that cannot be captured by hard outcomes, we should not hesitate to measure the QOL as the outcome of care.

Table 2. Disclosure of individual hospitals in their own websites

<table>
<thead>
<tr>
<th>Hospital</th>
<th>Patients</th>
<th>Follow-up rate</th>
<th>Stage stratification</th>
<th>Treatment period</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Surgical, all treated</td>
<td>99.1%</td>
<td>Yes</td>
<td>2005/1–2009/1</td>
</tr>
<tr>
<td>2</td>
<td>Surgically treated</td>
<td>NP</td>
<td>No</td>
<td>1993–2008/12</td>
</tr>
<tr>
<td>3</td>
<td>Surgically treated</td>
<td>NP</td>
<td>No</td>
<td>2005–06</td>
</tr>
<tr>
<td>4</td>
<td>Surgically treated</td>
<td>100.0%</td>
<td>Yes</td>
<td>1993–2002</td>
</tr>
<tr>
<td>5</td>
<td>Surgically treated</td>
<td>NP</td>
<td>Yes</td>
<td>2007</td>
</tr>
<tr>
<td>6</td>
<td>Surgically treated</td>
<td>NP</td>
<td>Yes</td>
<td>1994/1–2000/12</td>
</tr>
<tr>
<td>7</td>
<td>Surgically/ endoscopically removed</td>
<td>NP</td>
<td>Yes</td>
<td>1970/8–1984/1</td>
</tr>
<tr>
<td>8</td>
<td>All treated</td>
<td>NP</td>
<td>Yes</td>
<td>NP</td>
</tr>
<tr>
<td>9</td>
<td>Surgically treated</td>
<td>NP</td>
<td>Yes</td>
<td>2004/1–2008/12</td>
</tr>
<tr>
<td>10</td>
<td>Surgically treated</td>
<td>NP</td>
<td>Yes</td>
<td>Last 10 years</td>
</tr>
<tr>
<td>11</td>
<td>Surgically treated</td>
<td>100.0%</td>
<td>Yes</td>
<td>1991–2003</td>
</tr>
<tr>
<td>12</td>
<td>Registered to HBCR</td>
<td>96.7%</td>
<td>Yes</td>
<td>2002/6–2003</td>
</tr>
<tr>
<td>13</td>
<td>Surgically treated</td>
<td>96.3%</td>
<td>Yes</td>
<td>1993/1–2007/6</td>
</tr>
<tr>
<td>14</td>
<td>Surgically treated</td>
<td>NP</td>
<td>Yes</td>
<td>1993/1–2001/12</td>
</tr>
</tbody>
</table>

HBCR, hospital-based cancer registry; NP, not presented.

Table 1. Excerpt from the data quality standards defined by the Japanese Association of Clinical Cancer Centers for public disclosure of cancer survival data

---
<table>
<thead>
<tr>
<th>Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>The start of observation should be set at the date of diagnosis</td>
</tr>
<tr>
<td>The first diagnosed cancer should be considered for survival calculation</td>
</tr>
<tr>
<td>Hospital-based cancer registry should be the data source; a clinician database can be an alternative at the start</td>
</tr>
<tr>
<td>Vital status should be elicited for &gt;90% of the cases (aim 95%), or survival should not be calculated</td>
</tr>
<tr>
<td>Stage should be elicited for &gt;60% of the cases (aim 80%), or the survival should not be calculated</td>
</tr>
<tr>
<td>Vital status should be ascertained by referral to local government office data</td>
</tr>
<tr>
<td>Survival should be calculated with Kaplan–Meier estimation</td>
</tr>
<tr>
<td>Relative survival should be used to adjust for death from the other causes</td>
</tr>
</tbody>
</table>
---
Cancer care quality control in Japan through peer-review by clinical experts, or in an explicit way

"To overcome these limitations, the measurement of processes can be done by non-experts and has greater reliability (22). The implicit methods have much greater flexibility than the explicit methods because the care can be examined according to the reviewers’ expertise. However, the results can be unreliable without a structural guide for the review (21), and requires extensive involvement of experts, which is sometimes difficult to obtain because of their demanding schedule. In the method using explicit criteria, once the criteria are developed, data collection for the implementation can be done by non-experts and has greater reliability (22). It may not be an easy task to set the explicit criteria, but is probably a more realistic option than to have experts available for the review. Typically, explicit criteria define the standards of care with a description of target patients and the care processes that such patients should receive (6,19,23–25).

The proportion of target patients who actually receive the care specified is considered to represent the quality of care in that aspect. Therefore, such criteria are often called quality indicators (QIs, more specifically, process-of-care QIs, because outcome or structure can represent quality, too), and the proportion calculated are called quality scores. These QIs have some similarity with the recommendations in clinical practice guidelines in describing the standards of care, but the quality criteria must be more clearly defined for evaluation purposes. For example, one frequently used criterion, ‘post-surgical chemotherapy for Stage III colon cancer patients’ (26,27) defines the target patient as colon cancer patients who had surgical removal of the tumor and were pathological Stage III, and the care processes as chemotherapy within a certain time frame after surgery. Since each QI covers only an aspect of care for the specified patients, many QIs are needed to comprehensively examine the quality of care (6,24,25)."

The missing portion of the page would continue with more detailed information on the development of QIs, including a Japanese government-funded research project started in late 2006, and the process of developing QIs to measure the quality of care for breast, colorectal, liver, lung and stomach cancer and palliative care.
published on the project website (28). Examples of the QIs are listed in Table 3.

CONSIDERATION ON IMPLEMENTATION

For the research project, we used the primary source of information for QI implementation as patient medical records to minimize the restriction in the area for measurement. Since the designated cancer care hospitals are required to have cancer registrars to operate the hospital-based cancer registry, we expected that they would take on the task of record abstraction. We piloted the implementation of the QIs in cooperation with cancer registrars. In 2009–10, we conducted the pilot project in 18 hospitals nationwide to implement the QIs for stomach, colorectal, breast and lung cancers. We reviewed all newly treated cases in the participating facilities. The results found reasonable care in some areas and room for improvement in other areas (47–49).

Moreover, from the pilot studies, we found several challenges to be overcome in the establishment of a nationwide system. First, the task of medical record abstraction took up a large amount of time of the abstractors. The abstraction of one patient’s data could take 40 min to an hour. As of 2012, it is a requirement for the designated cancer care hospitals to hire at least one trained cancer registrar to manage the hospital-based cancer registry. If we are to measure the quality as part of the routine work of the registrars, we need more staff.

Secondly, the documentation is sometimes insufficient. It is particularly troublesome when information to determine the eligibility of patients for a QI is missing. For example, performance status was frequently missing thus diminishing our ability to determine the indication for chemotherapy. Furthermore, some findings were documented in order to support the treatment choice. For example, the level of pain may be documented only when the patient receives analgesic medications. If this happens, the proportion of patients whose pain was addressed and treated will be overestimated. On the other hand, lack of documentation can, of course, lead to poor quality scores. If the indicated care was performed but not documented, it will not be captured in the quality data, and thus, considered the same as not provided in the quality measures. The quality of documentation can substantially influence the measured quality. In a different perspective, however, the quality of documentation is certainly part of quality. Miscommunication between health providers is a major contributor of medical errors (50). Appropriate documentation enables information sharing among the team of health professionals, preventing errors and enhancing smooth collaboration. Documentation does not need to be over-emphasized, but should not be overlooked.

Another rare but interesting problem we found was that some records were written totally in English, not Japanese. Although English is a mandatory subject in the Japanese education system, we cannot expect that all medical record abstractors understand the documentation written in English. In such cases, medical record abstraction needs help from doctors or those who have proficiency in English. From the quality of care perspective, we do not know whether such documentation practices represent better or worse quality. It may represent the documenting physician’s ability to collect state-of-the-art medical information not only in Japanese but also in English, but it may weaken the level of information sharing among the local health-care team. This may need to be addressed by discussion with the clinical experts in the process of setting the standards for quality.

Table 3. Examples of QIs

<table>
<thead>
<tr>
<th>Numerators (care processes recommended)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gastric cancer patients with cT1N1–3 or cT2-4aN0-4 disease</td>
</tr>
<tr>
<td>Patients with colorectal cancer who underwent surgical resection and diagnosed pathologically with Stage III disease</td>
</tr>
<tr>
<td>Colorectal cancer patients who underwent surgery</td>
</tr>
<tr>
<td>Patients with Stage I to II breast cancer of 3 cm or less in diameter</td>
</tr>
<tr>
<td>Patients &lt;75 years old, who were PS0-1, diagnosed with Stages III–IV non-small-cell lung cancer, and received chemotherapy</td>
</tr>
<tr>
<td>Patients with hepatocellular carcinoma and liver damage class A, having 3 or less tumors of 3 cm or smaller in diameter</td>
</tr>
<tr>
<td>Patients started on opioid therapy on an outpatient basis</td>
</tr>
</tbody>
</table>

CONSIDERATION ON PROCESS–OUTCOME LINK

Theoretically, the process quality must be linked to outcomes. As the aim of medical care is to improve outcomes, the processes that do not improve the outcomes do not become the standard for high-quality care. Process-of-care QIs developed in prior studies were often examined for their link with improved outcomes. Some showed a positive relationship (35,51) and others did not (52). A study that compared the survival curves between those who received post-surgical drug therapy conforming to the St. Gallen recommendation and those who did not showed that the former had a better survival (51). In contrast, many
of the well-established QIs for chronic heart disease were not associated with the 60–90-day mortality after discharge (52).

However, we need to make distinction between the theoretical and observed link. It is reasonable to examine the process–outcome link in the real data, but several problems exist to overemphasize the process–outcome link observed, or the proposal to rank QIs based on the link to the outcomes. First, the outcomes targeted to be improved with the process vary across the QIs. For example, the QI that recommends prophylactic anti-emetic use along with high-risk chemotherapy aims to improve the QOL during chemotherapy, and the QI that recommends giving sufficient explanation to patients about their treatment focuses on patient centeredness. These QIs cannot be compared with QIs for care that improves survival. Furthermore, some care aims to improve short-term outcomes while some improves long-term outcomes (53). Secondly, diagnostic processes improve outcomes through enabling appropriate treatment, which is a more indirect effect on outcomes than the treatment processes. It may lead to the diagnostic processes showing a weaker link to the treatment processes. However, it does not mean that diagnosis is less important than treatment. Thirdly, even if we do not find a relationship between the process of care and outcomes in an observational study, the reasons could be anything other than the true effect of care processes, such as the sample size, and confounders other than the quality. In particular, when the care process proved to be effective in randomized controlled studies, a lack of relationship in the examination of process–outcome, which probably takes an observational design, does not necessarily refute the efficacy of care.

With all these limitations in mind, however, we agree that it is worth examining the process–outcome link. If appropriately interpreted, the results will enrich discussion by the experts in the revision or development of the QIs in future.

**Limitations of Process Measurement**

Compared with the outcomes, processes have their own limitations, too. First, the quality of care can be measured only in the area where standards exist and are contained in the QIs. Innovative approaches to advance medical care or excellent surgical skills for which no standard exists cannot be measured. Such aspects of care are expected to be captured in the outcomes measurement. Secondly, the content of processes needs expert knowledge and the relative importance of each can sometimes be uncertain. It leads to uncertainty in creating a valid summary score integrating multiple QIs. Finally, the criteria can change along with advances in medical knowledge. A prior study showed that half of clinical practice guidelines are out of date in 5.8 years (54). Process of care QIs need periodic updates, and once the QI changes, we cannot trace trends in quality. Outcome measures are more stable in this sense. For quality measurements to be useful and reliable, a correct balance of processes and outcomes is necessary.

**Current Activities for the Process Quality Monitoring System**

In December 2011, the Association of Prefectural Designated Cancer Care Hospitals started the Cancer Registry Chapter, which sponsors the QI activities. These activities are intended to use hospital-based cancer registry data linked with health insurance claims. The hospital-based cancer registries contain patient characteristics and tumor characteristics, including cancer type, histopathology and TNM stage, and the health insurance claims data contain the medical services provided in the facility. The hospital-based cancer registries can be used to create a list of eligible patients because these data are routinely submitted to the National Cancer Center every year. The greatest advantage of these data sets is that they are all electronically available.

A possible limitation in the use of these electronic data is that we cannot capture the care provided in other facilities. For example, as mentioned above, a QI states that Stage III colorectal cancer should receive chemotherapy after surgery. If patients were referred to another hospital after surgery and received chemotherapy in the other hospital, this chemotherapy cannot be captured in the claims data of the original facility. Such collaboration in care between hospitals is now encouraged for the efficient use of resources. We will need to examine the extent that this referral practice influences the quality measurement in the course of our quality measurement activities.

Another limitation is that we cannot take into account the reasons for care inconsistent with the standards recommended in the QI. The recommended care may not be provided because of patient preference or patient physical conditions such as age and comorbidities. The electronic data does not include information on such clinical judgment. From the report on the Quality and Outcomes Framework in the UK, these exceptions existed in ~5% of cases (55). This figure may be different in cancer care and will also differ across the types of care and patient populations. A separate study will need to address the level of exceptions.

Given the two limitations, caution will be needed in considering the results as ‘quality of care’. One additional step will be necessary to ascertain the care provision in other facilities and reasons for exceptional treatment. Nonetheless, identifying the cases who failed to receive the recommended care and examining the reasons will be part of the quality improvement processes. The benefit can become even greater if the reasons for exceptions are shared widely and the ideal care for such cases is discussed. It will reveal areas where controversy and variation exist for future clinical research, which should advance the knowledge of patient management. After all, this may be the most practical way of constructing an ongoing monitoring system of quality of cancer care in Japan.
CONCLUSIONS

We presented an overview of current activities for measuring the quality of care in Japan. In the era of population aging and rising health-care costs, ensuring the quality of care is of greater importance ever before. The quality measurement is not a simple task. There are a myriad of considerations in preparing, using and interpreting the measurements. Nevertheless, to ensure and improve the quality of care, we must start by measuring. Although the perfect measurement of quality is difficult, measurement does not need to be perfect to enable improvement. We simply need the data to act on. A key to improvement is not to blame someone for the problems found or simply to compete against each other, but for every player to work collaboratively to solve the problems. Whether we should make the results public may be a concern among those measured, but it is a secondary issue. The research findings to date show patients do not frequently use the quality measurement for their chosen hospital (56, 57), but transparency is a value in itself unless it causes unintended consequences, such as health providers gaming with the measurement (58). We should continue efforts to measure the quality, discuss improvements in both measurement and care and research the answer to questions arising during the processes.

FUNDING

The research project for the development of process-of-care QIs received funding by a grant-in-aid from the Ministry of Health, Labor and Welfare, Japan.

Conflict of interest statement

None declared.

References


