Note 5: Outcome Measures: Practices and Methodologies

- Prof. dr. Fred van Eenennaam and Kim Brüheim, MSc.

Over treatment, under treatment, over diagnosis, under diagnosis, uncontrolled costs and budgets, medical treatment errors and wrongly placed incentives have been reported across health systems in the developed world. The framework, concepts, practices; theories and tools of Value Based Health Care Delivery (VBHC) coined by Harvard's Prof Porter, as indicated in his famous article of the New England Journal of Medicine (Porter, M.E. 2010), have and are being developed internationally by amongst others Prof. Porters and Prof. Teisberg's group at Harvard University and Business School as a central node reference.

The issues, faced across different advanced health care systems, are addressed from a decision makers and clinical management perspective. The aim of this note is to share key practices and core methodological experiences in applying the outcome measures component of VBHC.

This note focuses on the part of:

1. Patient relevant medical outcome measures (outcome measures) as a core component for decision makers and clinical management.
2. Organizational and communication challenges on outcomes
3. Decision making methodologies on outcome measures selection, patient initial conditions, comparisons, learning and incentivizing from a health management and leadership perspective.

1. Outcome Measures within VBHC

The patient value view must be central, when making responsible decisions by decision makers and clinical management, within the health system and health delivery organizations. More precisely patient value is patient outcome divided by costs. Although most decision makers, practitioners, health scholars, policy makers and economists concur with this notion, in practice often the opposite views have been implemented in health care systems.

Clinical management, doctors, nurses and other (para)medical teams need to contemplate their (non)interventions across the whole medical condition with the patient outcome and costs. However, too often, they are involved in: non-related balance score cards, specific lean units, allocated budgets or arbitrary costs of business units, and reporting endless data on which decision nor their own nor of others can be based.

The management of the health care organization needs to be able to make decisions on what to do and not to do, the reimbursements policies and negotiations on the value that is delivered to patients, and the costs and the risks management of the organization. While CEOs of hospitals embarking on VBHC initiatives have reported difficulties in lacking adequate information while being flooded with data on processes, structural (Dimick J.B. et al 2004), “quality, morbidity data”, score cards, confusing costs and tariffs/DRG schemes, capital expenditures, detailed clinical research outcomes, annual reporting data etc.

Payers, as in health insurance firms, patients, governments, oversight agencies or others, need to decide: when to receive treatment or not, at which place based on their medical condition(s)/co morbidities, and at what likely outcomes, costs and risks. Accepting that doctors and hospitals cannot be everything to everybody, no patient exactly fits a protocol and no patient is 100% unique. From this medical perspective, it is challenging to search for right diagnosis and right treatment (Christensen, C. M. 2009).

The health sector has long stated that a positive form of (regulated) competition like in the rest of society and the economy has existed, is not achievable. A doctor, team, health organization stating (and verified) that its patient value or offer is best for patients with XYZ seeking patient outcome ABC with a related cost of ZZZ, has a positive patient value statement. It allows for specialization and concentration for diagnosis, treatment and care at better outcomes per unit costs, since health providers compete. This positive force of competition can only be achieved when outcome measures are known and are incorporated at the decision making moments of the key actors.

2. Outcome Measures in the VBHC Practice

Two categories of challenges are faced from a decision making and managerial perspective when outcome measures are implemented and used in the VBHC practice; First the organizational legacy challenge, and secondly the communication challenge.

2.1. Organizational legacies:

Two categories of challenges are faced from a decision making and managerial perspective when outcome measures are implemented and used in the VBHC practice; First the organizational legacy challenge, and secondly the communication challenge.

- The main issue is that medical and non-medical staffs have become tired of the time consuming tasks of collecting of all kinds of data. Many of the process, structure, cost, compliance and patient data collected are often rightfully seen as wasteful activities. The relevance and use of the collected data is unclear, not related to patients or the medical or managerial tasks. It is unclear, what is done with the data or how it helps improve or take decisions. The most positive "feedback of the data" is often in the form of "spiegel data" with a perceived clear cost reduction aim. The questions asked is: why do you use twice the medicine, sacks of blood, or other "wastes" of materials unrelated to clinical or patient work. The question should be: why do you not use twice as much of the medicine or diagnostics in the interest of patient and the saving of costs later on.

  The "tiredness" and "motivation" to collect and improve is easily overcome, when the focus is set on outcome for patients. Since these outcomes relates with a systematic approach and a clear link to the decisions that must be made by those involved. Our VBHC projects and other VBHC projects around the world are starting to show these positive effects. A positive side effect is often a pruning of existing indicators.

- The second organizational legacy is the departmentalization along medical specialties/interventions and the lack of using the latest medical insights across medical conditions and related outcomes. Typically the training and organizational departmentalization has a focus on one's own specialty/medical intervention and the budget of the current unit. Consequently "outcomes" are reported on typical process or structure indicators of the individual interventions or organizational unit. A process measure may be for example, the use of 2 liters of blood and 7.5 minutes for the intervention. This process measure may or may not be related to the overall outcome measure, it is certainly not useful for decision making for anyone outside the direct operational unit. Unfortunately, clinical studies were process and structure indicators clearly link to outcome measures medically and statistically are limitedly availably.

  Outcome measures which are to measure results to be achieved across units or departments or "zorgketens" are hard to achieve, although efforts on team work and integration are underway and some results have been reported. Examples are multifold were for instance patients may benefit from a certain more expensive stent placement by

unit one, which will result in better overall outcomes will however increase costs in unit 1 and lower costs in unit 5. In practice to determine across departments outcome measures and consequently improvements in across department behaviors even at the expense of some departments depends heavily on the informal notions of team work across departments and are hindered by the organizational structures[1].

These first two legacies are not hard to overcome when the relevant medical and clinical management groups are gathered taking the outcomes together with outside validation and involvement. This is based on experiences from the authors and others around the world.

- The third organizational legacy, the lagging ICT systems and data management are a much harder to overcome, and demand a 3 to 5 year time horizon. The requirements on information of outcomes are different than the current demands on ICT systems. Data management and the data warehouse needs adjustment when the structure and management has used another structure and point of view. Although the required data for outcomes is not very complex or difficult, other demands on ICT systems, sometimes outdated ICT architectures and limited ICT human capacity requires some time before the relevant outcome data can be automatically captured.

To have a reasonable set of relevant data for decision making, the systems, data and reporting organization requires a planning for the next years. Generally on reporting the information it is practical to start with a small and highly meaningful set and building forward while more in depth data will be available over time. In practice an extra compatible system is usual set up, were the medical records of the past are recoded to get the required data in place. The sooner this data can be incorporated in the system, the better. This typically happens when a new system, upgrade and/or data architecture effort is planned. The building up of large data ware houses based upon existing data to do some pure data analyses (data fishing exercises) has limited usability in detecting some potential for partial improvements or detecting some correlations. Since the data does not contain outcomes measures the relevance for decision making is negligible or even negative when it adds to the data flood of decision makers.

2.2. Communication challenges:

Communicating the results to the public/press is a major concern for those involved. Some professionals, who fear a blaming game, will only communicate positive news or have tendency to report only very nuanced results. The real advice is to get the communication and PR departments involved on how the results can best be fairly communicated. Media and crisis communication trainings have become a standard practice for higher ranking professionals. Social media and the internet are enlarging this challenge as well as the opportunity.

Another important aspect is that communicating results by different decision makers results in different styles and levels of reporting. The main idea on outcome measures is that all parties

[1] The so called lean and process improvement efforts sometimes have been reported to have adverse effects. Most notably were only current practices in existing departments/interventions have been performed more efficiently without a link to outcomes or the latest clinical management insights and practices.

involved can easily understand the most relevant 3 to 7 measures on a specific medical condition. All parties have then a similar starting point.

- For payers, patients, government agencies that level of reporting should normally be enough for the decisions to consider a hospital or medical team for a diagnoses and/or treatment. When a score rises often major concerns or major innovations take place in a certain medical condition. As a result the focus can be put a level deeper on that specific medical conditions.
- Color coding and other coding methods have been used for many years in industry. Hospital management is often helped as well with an overall score of a medical condition mainly reducing the 3 to 7 outcomes to one score. A warning will pop up if one of the outcome measures needs attention, of an individual medical condition, to allow for a deeper level analyses and decision making. This cascading of outcome measures has been commonly used on financial information and can easily be applied on the outcome measures once those involved trust the system and let it be audited regularly from the outside (not by audit firms).
- For clinical management, the physicians, nurses and others involved the outcome measures are just the start of their work. They need further refined analyses on the patient initial conditions, process and structure indicators and VBHC tools to analyze, learn, improve and check their processes, tasks, systems, technologies etc. Outcome measures relate back to the detailed practices in the medical conditions and helps medical and non-medical staffs (as well as industry) focus on improvements that create patient value.

From a decision making stand point, it is wise to report the outcome measures raw, meaning without any patient initial condition/case mix adjustments first. The main reason is that the strongest claim for medical professionals is to score well without the need for adjustments. If reporting of patient initial conditions are thought to be fair than there is a preferred order of reporting.

The preferred order of reporting outcome measures is:

1. Always report the raw unadjusted data
2. Report in comparison with similar hospitals and medical centers.
3. Report only the variance in patient initial conditions and keep the overall outcome measures
4. Report cross tabs with (main) initial conditions and outcome measures
5. Use statistical methods to correct for case mixes[2]

[2] See (Bosch van den, W., 2011; Bosch van den, W.F. et al. 2010; Bosch van den, W.F. et al. 2009; Geelkerken, R.H. 2008; Linford, R., Pronovost, P., 2010) for difficulties with morbidity corrected measures. The assumed normality in most statistical techniques are often challenged, while efforts to come with one general correction factor of case mixes or one algorithm have met severe criticism. Until now these methods are not widely accepted and communicable, consequently their relevance for decision making is limited on the level of outcome measures. These statistical techniques are undisputed at the level of clinical research and simple, standard, low variance decision making and are frequently used by the authors in operations management questions or analyses.

The more complex and nuanced the reporting becomes the less useful the information is for the decision making by payers, policy makers and hospital management. Individual patients will use the general information for their first selection and will need individual and family counseling along their diagnosis and treatment trajectory.

There is a tendency for government agencies and payers to ask for more data when data is available. The effect is that reporting gets more attention and enable medical professionals or hospitals with a good system compared to those that do not have that. This governance paradox of having a heavier oversight burden when you do a good job than when you do a poor, is dangerous and asks for restraint from oversights bodies.

3. Determining Outcome Measures, Initial Conditions and Methodologies

The methodologies used to determine outcome measures (and VBHC) are based on decision making sciences. On the first level, where the decisions are simple, standardized, not having much variances, and the data is sufficiently available, simple statistical techniques like multiple regressions etcetera can be used.

On the progressing next levels, when decisions become less simple, standardized, variances increases, and data is a problem, we use all possible methodologies from level one. However that is not enough to help decision makers take decisions, so we use a set of extra methodologies namely: decision making under uncertainty, decision making under imperfect information, single and double loop learning methodologies, systems theories and system dynamics.

Interestingly, the methodologies often referred to as 'evidence base' or 'level one', are a part of the methodologies which are not appropriate as basis for a decision. In reality decision makers have very limited benefits from the analyses of methodologies of level one. They are faced with complexities of running organizations such as hospitals. They systematically address the issue that the methodologies which consistently requires learning and improvement within and outside the organization. The outcome measures methodology fits that decision making methodologies tradition.

3.1. Selection of outcome measures

Assuming medical condition & care delivery value chain (see note 4).

Based on the analyses of amongst others Prof. Porters and Prof. Teisberg's group at Harvard University and Business School (Porter, M.E. 2010; Porter, M. E., & Teisberg, E. O. 2006; Porter, M. E. 1996.) the following methodology is developed and employed to determine outcome measures which allow decision making in VBHC.

Aim is to select between 3 and 7 measures[3] which are most important for the patient, based on Porters Outcome Measures Hierarchy.

[3] In practice another specialist group will add 2 to 3 which are typical for them and shows specialization effect.

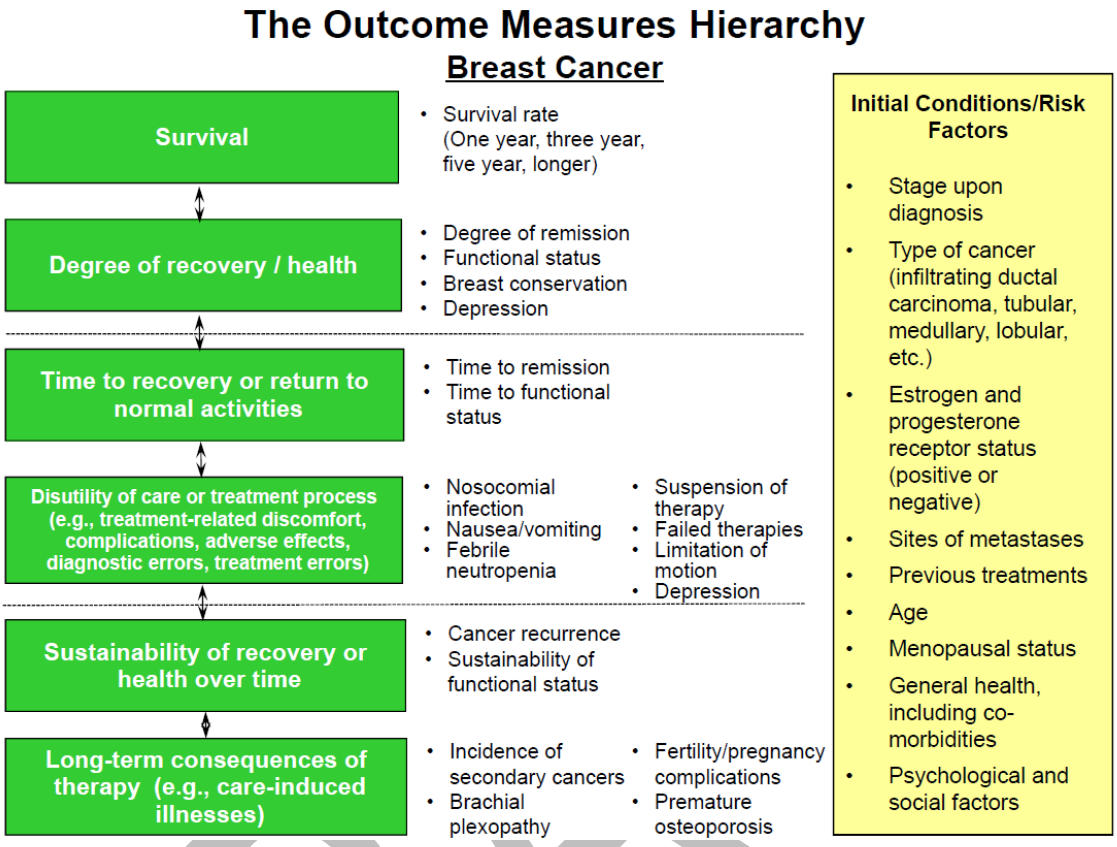

Figure 1 Outcome Measures Hierarchy: Example of Breast Cancer (Porter, M.E. 2012)

- The latest clinical research methodologies are used to create a list of available outcome measures. A broad list of potential outcome indicators, practiced by outside agencies, professionals, patients, World Health Organization etc., is set up.

There are already outcome measures known, used, and validated by international institutions like the Cleveland Clinic, USA. Learning from each other, and using experiences, expertise and knowledge results in the installation of new learning culture in the organization. The best practices of international and renowned health organizations are guidelines for your own organization.

In line with and based on the medical condition and the complete care delivery value chain, where all value-adding activities are included, the best patient relevant medical outcome measures are chosen. The broad list will be added with new outcome measures. In practice, due to defining the medical condition, and care delivery value chain, the value adding activities become visible. Professionals will realize what the most important aspects are in patients perspective. These new insights lead to possible new outcome measures. After completing the list, the selection to the 3 to 7 measures is based on four selection phases.

- The first selection phase is based primarily on patient value. This is the most important aspect, and requires intensive workshops and extensive discussions for optimization.

We always make sure that we have the most relevant outside validators in to avoid group think.

- Secondly, the user value for clinical management and decision makers will be taken into account. Can we make decisions based on these outcomes, and what will be the impact of these decisions?
- Thirdly, the practicality measured in feasibility of data collection and quality of the outcome definition will play the last and minor role in the decision from the distillation of circa 100 outcome measures to 3 to 7 measures.

The selection procedure is made by an expert team and supported and validated by international academic experts. These experts can be divided in two groups, a medical and a methodology group. The medical group consists of renowned and international experts on the specific medical conditions. The methodology group consists of experts in data analysis, validation and the principles of VBHC.

The complete process of the establishment of the selection of old and new outcome measures, results in a shift in mindset. Professionals are more aligned with the expertise and value adding activities of their colleagues in within other specialties. Their focus switches to a share way of working towards better patient values.

3.2. Patient initial conditions/case mixes

Patient initial conditions, or in Dutch the case mix, presents a risk adjustment for the selected outcome measures. (See 'communication section' on sensitivity in communication of results for medical doctors and hospitals.)

- With the same procedures, and methodologies a broad list is created with relevant patient initial conditions. Also by the definition of this list, the vision of clinical management and clinical research is the leading guideline.
- The first selection is based on the patient initial conditions, which are known to explain by literature to have the most variance.
- The remaining patient initial conditions are validated by experts, in line with the selection of outcome measures.
- Based on experiences, two extra patient initial conditions are added by doctors/professionals because of their clinical view, since clinical research may not have corroborated clinical views that are widely held.

We must accept that given the complex nature of medical problems in organisms, the existence of co morbidity, that the variance explained evidence base is generally limited. In the future when more and more personalized diagnostics and treatments will be developed, initial conditions are likely to be more meaningful and available. Today co morbidity, age, diabetes, smoking, exercising or not, are usually the core group of patient initial conditions.

3.3. Comparisons, learning and incentives

When the outcome measures and the patient initial conditions are set, the question becomes what to do with the outcome measures.

- First of all, the track record over time of the treatment of the patients in the medical condition can be compared and progress (learning) observed. Often the trace back of data for some years, using medical patient records, gives a starting point.
- Secondly comparing with institutions that are among the leaders in the world on diagnosing and treating the same medical condition is useful. They will have a similar core group of outcomes measures at some point. Within the Netherlands/Europe, other health systems can be added easily over time.
- Then seeking the most relevant reference group, in an international context. A preferred group is a mix of institutions with: different volumes, different health or outcome effects, different costs structures, different innovative profiles, and different practices.

The selection of outcome measures, the i patient initial conditions and its usage will become more and more relevant in decision making over time for decision makers and will be constantly improved. This is referred to as a single loop learning process itself. Where patient value is increasing over time, as defined by patient outcomes divided by cost, a double loop learning process is implemented from the VBHC aim of continuous improvement and positive (regulated) competition.

When enough data is available one statistical analyze can be conducted, refining and pruning the use of different outcome measures and patient initial conditions.

For the Dutch health care system a key point is to determine how the benefits of more patient value will be divided among the stakeholders. Achievements for patient value will come from medical staff and non-medical staff; they should be reasonably incentivized to do so with an upward and downward risk sharing.

4. Conclusion

Note 5 focuses on the part of patient relevant medical outcome measures (outcome measures) as a core component for decision makers and clinical management applying Value Based Health Care (VBHC). The aim of this note is to share key practices and core methodological experiences in applying the outcome measures component of VBHC. Key practices on outcomes organizational and communication challenges and decision making methodologies on outcome measures selection, patient initial conditions, comparisons, learning and incentivizing from a health management and leadership perspective.

Outcome measures should be viewed in conjunction with cost to allow for a patient value assessment or cost neutrality must be assumed.

Outcome measures are reasonably easy to establish but will take some time to provide the benefits for decision makers like payer, patients, policy makers, doctors, hospital management and other stakeholders. Across the world VBHC and outcome measure are beginning to show a way forward out of the almost universal problems of health care systems on over treatment, under treatment, over diagnosis, under diagnosis, uncontrolled costs and budgets, medical treatment errors, and wrongly placed incentives.

Partial Literature List

Allison JJ, et al. Kiefe CI, "Improving quality improvement using achievable benchmarks for physician feedback: a randomized controlled trial.," *JAMA*, vol. 285, pp. 2871-2879, 2001.Christensen, C. M., Grossman, J. H., & Hwang, J. (2009).*The innovator's prescription, a disruptive solution for health care*. McGraw-Hill; New York.

Bosch van den, W.F., (2011) De HSMR beproefd – Aard en invloed van meetfouten bij het bepalen van het gestandaardiseerde ziekenhuissterftecijfer. *Academisch Proefschrift*., GVO Ponsen&Looijen, Ede, Nederland

Bosch van den, W.F. et al. (2010) Variatie in coderingpatientgegeens beinvloed gestandaardiseerd ziekenhuissterftecijfer (HSMR). Nederlands Tijdschrijft voor de Geneeskunde 154;A1189.

Bosch van den, W.F. et al. (2009) Schommelende Sterftecijfers, HSMR nog geen betrouwbare maat voor zorgkwaliteit. Medisch Contact, 64:1344-8.

Dimick, J.B., Brikmeyer N.J.O., Brikmeyer J.D. (2004). Measuring the quality of surgical care: structure, process or outcomes? *Journal of the American College of Surgeons*, vol. 198, no.4, pp. 626-632

Geelkerken, R.H. (2008) Een onrijp instrument – Sterfte cijfer niet geschikt als maat voor ziekenhuiskwaliteit. Medisch Contact nr. 09. 370-374.

Hunt S.S, Maerki S.C. Luft H.S., (1987) *The volume-outcome relationship: practice-makes-perfect or selective-referral patterns?,* Health Serv Res , vol. 22, no. 2, pp. 157-82.

Institute of Medicine, Crossing the Quality Chasm (2001). A New Health System for the 21st Century. Committee of Quality of Care in America. Washington DC: National Academy Press.

Janlov, N. (2007). Swedish Health Care Performance – Quantity versus Quality, Department of Economics, Lund University

Linford, R., Pronovost, P., (2010) Using hospital mortality rates to judge hospital performance: a bad idea that just don't go away. BMJ 340:c2066.

Meredith, B., et al. "Pay for Performance: A Decision Guide for Purchasers," University of California, San Francisco, Harvard School of Public Health,.

Porter, M.E. (2012) Value Based Healthcare Delivery – Core Concepts. Presentation Harvard Business School, January 4.

Porter, M.E. (2010) What is value in healthcare? *The New England Journal of Medicine*, 363;26

Porter, M.E. (2007) Creating a High-Value Health Care System: Implications for Sweden. Harvard Business School.

Porter, M.E., & Teisberg, E. O. (2006). *Redefining health care; creating positive-sum competition to deliver value.* Boston, MA: Harvard Business Press.

Porter, M.E. (1996). What is strategy? *Harvard Business Review,* November-December; 59-78.

Rapp, M.T. (2009). Outcome Measures and Value Based Purchasing. *Presentation given at AHRQ Annual Conference*

Romney, A.C., Bohmer, R., (2009) Performance Measurement at Intermountain Healthcare, Harvard Business School.

USA Department of Health and Human Services – Food and Drug Administration (FDA), (2009). Guidance for Industry; Patient-Reported Outcome Measures: Use in Medical Producut Development to Support Labeling Claims.

Prof. dr. Fred van Eenennaam and Kim Brüheim, MSc. partner and consultant of The Decision Group, have an extensive experience in the principles of Value Based Healthcare. At this moment, The Decision Group has experiences in the specialisms and medical conditions with the highest prevalence (Cardioloy-AMI, CAD; Oncology-Breast; Orthopedics-Hip Dysplasia; Neurology-Alzheimer's Disease; Internal Medicine-Diabetes; Gynacology-Obstetrics; and General Clinical Management. Visit for more information www.thedecisiongroup.nl or www.thedecisioninstitute.org or contact k.bruheim@thedecisiongroup.nl