Resource stewardship projects during postgraduate training

A toolkit for faculty supervising resident research and quality improvement projects

Part of the CanMEDS Resource Stewardship Curriculum Toolkit Series

Editor
Jerome Leis, MD, MSc (QI/PS), FRCPC
Division of Infectious Diseases
Sunnybrook Health Sciences Centre
Department of Medicine
University of Toronto
Toronto, Ontario

Series editor
Chris Hillis, MD, MSc (QI/PS), FRCPC
Department of Oncology
McMaster University
Hamilton, Ontario

The Royal College of Physicians and Surgeons of Canada
774 Echo Drive
Ottawa, Ontario K1S 5N8 Canada
Telephone: 613-730-8177
Toll-Free: 1-800-668-3740
Fax: 613-730-3707
Website: royalcollege.ca/canmeds
E-mail: canmeds@royalcollege.ca

How to reference this document:
Hillis C and J Leis, editors. 2017 *Resource stewardship projects during postgraduate training. A toolkit for faculty supervising resident research and quality improvement projects*. Part of *The CanMEDS Resource Stewardship Curriculum Toolkit Series.* Ottawa: The Royal College of Physicians and Surgeons of Canada.

CONTENTS

PREAMBLE

Why undertake a resource stewardship project?

The inclusion of resource stewardship within the Leader Role of the CanMEDS 2015 Framework[1] reflects a growing emphasis on the appropriate use of health care resources. Resource stewardship is a specific competency that requires training and experience, and leading a resource stewardship project is an opportunity for residents to demonstrate competence in this area while completing a scholarly project. Developing skills in resource stewardship is essential to physicians' leadership in the development of an integrated health care system that promotes high-value care.

Who are the intended users of this toolkit?

During residency, trainees are expected to undertake a scholarly project related to research or quality improvement (QI). Resource stewardship projects are one type of initiative that can be undertaken to fulfill this requirement. Although many resident-led resource stewardship projects have already been successfully completed, faculty experience in supervising these studies is variable. Meanwhile, with the introduction of CanMEDS 2015, we can expect that more faculty will be asked to supervise resource stewardship projects.

This toolkit is intended to assist faculty with limited experience in the planning, supervision and assessment of resource stewardship projects. It assumes that the user has some familiarity with QI methods and is looking for additional skills relevant to resource stewardship. For faculty who would like a refresher on how to conduct QI projects, the following resources are available:

- *Quality Improvement Guide*[2]
- *Teaching Quality Improvement in Residency Education*[3]

For those who would like a comprehensive, step-by-step guide to designing, conducting and reporting on health sciences research, we suggest:

- *The Research Guide: A Primer for Residents, Other Health Care Trainees, and Practitioners*[4]

The user of this toolkit is encouraged to begin with the "Getting started" section, which outlines the five main types of resource stewardship project. In the "Key steps" section, each project type is described in detail in five independent modules. An assessment tool is provided at the end of the toolkit, along with tips for supervising a successful resource stewardship project.

GETTING STARTED

What are resource stewardship projects?

The purpose of a resource stewardship project is to identify, understand or mitigate the overuse of diagnostic tests, treatments or procedures in health care. These projects overlap significantly with QI initiatives and may focus on any of the six domains of quality:[5] effectiveness, efficiency, patient safety, equity, timeliness and patient-centredness. The scope of a resource stewardship project may be limited to identifying and describing an area of overuse; understanding the key drivers of overuse; testing a change to address resource overuse; implementing a system change; or evaluating the sustainability or potential harms associated with a change. The project may be undertaken locally, in one health care institution or patient area, or may extend regionally or nationally. Whatever their scope, resource stewardship projects share common features – described below – that all faculty should be comfortable explaining to trainees.

What are the different types of resource stewardship project?

There is no one-size-fits-all design for a resource stewardship project, since different projects may tackle different aspects of the same resource stewardship problem. Despite variability in scope or focus, resource stewardship projects always begin with one or more aspects of a resource stewardship problem or question. Clearly defining the facet of the problem that the study aims to address will help to constrain the scope of the project and increase the chances of completion. A trainee who completes one aspect of a resource stewardship project might have an opportunity to move to the next phase in a subsequent stage of his or her training. Alternatively, the project might be handed off to a different trainee.

For the purposes of this toolkit, resource stewardship projects are classified into the following five types, according to the nature of the problem or question and the intended outcome:

1. ***Identifying or confirming a resource stewardship problem.*** These projects are focused on documenting resource overuse. They may involve performing a prospective or retrospective audit of local practice in order to document patterns of testing and treatment, with the aim of determining their appropriateness. The prospective audit of urinalysis orders in the emergency department described in Case Example 1.2 is one example of this kind of project.
2. ***Developing measures for resource utilization.*** These projects focus on standardizing and validating methods for measuring resource utilization, whether by proposing standard definitions regarding appropriateness or by validating new tools to measure test utilization. The validation of an automated documentation system as a means of determining rates of urinary catheter use in a hospital setting is one example (see Case Example 2.2).
3. ***Understanding the patterns or key drivers of overuse.*** These projects take a closer look at practices that are already known to represent overuse. The primary goal is to characterize patterns of behaviour or other key factors that contribute to the overuse. This usually involves a retrospective analysis of a clinical practice to identify predictors of overuse. For example, a root-cause analysis of sedative use among elderly patients in hospital could be undertaken to identify factors that could be addressed to remediate the problem (see Case Example 3.2).

4. ***Undertaking a test of change to reduce overuse.*** These projects usually involve a pilot or proof-of-concept trial to demonstrate the impact of a direct intervention on a resource utilization problem. The selected intervention should be based on a well-founded hypothesis regarding the key driver of overuse and use an iterative process to make refinements as needed. An example of a system change might be the introduction of an electronic prompt to reassess standing bloodwork orders on a daily basis, with the aim of reducing unnecessary testing.
5. ***Evaluating the safety and sustainability of a new process of care.*** These projects involve larger-scale, longitudinal evaluations of a new process to evaluate whether it has achieved a sustained reduction in overuse without undesirable consequences. In these projects, the intervention has already have been developed and implemented successfully using iterative cycles of change, but there is a need to evaluate the impact of broader implementation. An example might be a 12-month evaluation after the elimination of screening urine cultures before joint replacement surgery to assess the impact of the change on laboratory costs, antibiotic use and prosthetic joint infection rates (see Case Example 4.2).

When is research ethics board approval necessary for a resource stewardship project?

The need for research ethics board (REB) approval for a resource stewardship project will depend on whether the primary focus of the project is QI, research or both. The following examples should help to illustrate when REB approval is required with respect to the five project categories described above, but faculty and residents should always consult with their institutional REB before undertaking a project.

- ***Identifying or confirming a resource stewardship problem.*** Conducting audits of utilization rates of tests and procedures is an expectation of physician professional practice and generally does not require REB approval. One exception would be an audit involving a broader patient population than would normally be seen by the faculty member or resident. For instance, if a resident were to audit the rates of pre-operative electrocardiograms (ECGs) in his or her own practice area, this would be a routine process improvement activity and would not require REB approval. Conversely, if the project required ECG ordering data to be obtained from another clinic, the REB might require a review of the study, since it would extend beyond the patients for whom the resident normally provides care.
- ***Developing measures for resource utilization.*** Projects focused on validating the measurement of resource utilization may or may not require REB approval, depending on whether the validation is being conducted locally for QI purposes, or more widely to create new knowledge about measurement across organizations.
- ***Understanding the patterns or key drivers of overuse.*** Projects that seek to characterize the larger pattern and drivers of overuse by conducting a formal retrospective analysis often require REB review to confirm whether individual patient consent for the use of their medical data can be waived.

- ***Undertaking a test of change to reduce overuse.*** Projects that undertake tests of change to reduce overuse can take many forms, ranging from iterative "plan–do–study–act" projects to formal, quasi-experimental evaluations. These projects may require REB approval, depending on the nature of the evaluation and the degree of potential risk the change is considered to have for patients. The ARECCI Ethics Screening Tool devised by Alberta Innovates[6] may be helpful in determining whether REB approval is required, but trainees undertaking such projects should also check with their institutional REB.
- ***Evaluating the safety and sustainability of a new process of care.*** Because these projects typically involve a research question that aims to create new knowledge regarding the effects and outcomes of new systems of care, REB approval is generally required. Even when the initial intervention was implemented for QI purposes, the evaluation aimed at creating new knowledge may require REB approval.

If a report on a resource stewardship project is submitted for publication, scientific journals will generally require that research ethics be specifically addressed, by stating either that REB approval was obtained or that the need for REB review was waived because the study was deemed to be a QI project. Since most residents will hope to publish their work, REB review requirements should be discussed before any data collection begins.

Who should be on the project team?

As in other types of QI projects, engaging all key stakeholders early in the planning phase of a resource stewardship project is critical. Although trainees may lead the project, faculty should ensure that the project team includes representatives from key stakeholder groups. In some situations, interdisciplinary supervision of the trainee can enhance the quality of the project by introducing a broader range of perspectives and experiences. For example, including a patient representative on the team, where possible, can help to promote patient-centredness. Moreover, having the right team composition will significantly increase the project's chance of acceptance by patients and other key stakeholders and, ultimately, its success.

KEY STEPS IN CONDUCTING A RESOURCE STEWARDSHIP PROJECT

In this section, the five types of resource stewardship project are described in further detail, along with key steps for their completion. The user of this toolkit is encouraged to select the module that applies to the type of resource stewardship project that he or she will be supervising.

Module One: Projects designed to identify or confirm the existence of a resource stewardship problem

Finding a topic

When a resident has a good topic in mind, the role of supervising faculty at the beginning of a project can be quite straightforward. In many cases, though, residents will look to faculty for guidance in choosing a topic. There is no shortage of resource overutilization problems in health care. Remind your residents that these can be identified from many sources, including personal experience. For example, a resource utilization issue observed in a particular clinical case can pique a trainee's interest and encourage him or her to explore whether it is symptomatic of a larger problem that might potentially be improved through a resource stewardship intervention. Other sources of information or inspiration include the lists of questionable tests and procedures compiled by national specialty societies for the Choosing Wisely campaign.[7] These lists specify tests, treatments or procedures that have no evidence of benefit or are potentially harmful to patients (see Table 1.1 for examples). Local institutional data may also point to areas of overuse and provide an incentive to improve practice.

Table 1.1
Examples of Choosing Wisely recommendations that may inspire resource stewardship projects

Don't order routine chest radiographs for critically ill patients, except to answer a specific clinical question.
Don't do imaging for lower-back pain unless red flags are present.
Don't routinely perform preoperative testing (such as chest X-rays, echocardiograms, or cardiac stress tests) for patients undergoing low risk surgeries.
Don't use benzodiazepines and other sedative-hypnotics in older adults as first choice for insomnia, agitation or delirium.
Don't recommend percutaneous feeding tubes in patients with advanced dementia; instead offer oral feeding.

Sample size calculation

Residents may need some guidance in establishing an appropriate sample size. Remind them that before they conduct a practice audit they will need to calculate the minimum sample size necessary to confirm or disprove the existence of resource overuse (see Case Example 1.1). This sample size will differ according to whether the objective is simply to confirm the *presence* of overuse, or to provide a precise estimate of the *degree* of overuse. The perception stemming from traditional research projects is that larger sample sizes should be always be used to enhance the accuracy of an estimate. However, in resource stewardship projects, although one wants to arrive at an *accurate* conclusion, getting a *precise* estimate is not always the immediate goal.

Sometimes the first step in a project is simply to confirm whether a perception of overuse is correct. A research question framed in terms of whether one's practice is meeting a reasonable and accepted target can sometimes be answered with a small sample, as a narrative review by Etchells and colleagues explains.[8] Using probability calculations, they provide a table that can be used both to determine whether the results of an audit with a given sample size is statistically significant and to plan the sample size for an audit (see Table 1.2 below). For example, if a completed audit shows 50% compliance when 80% is the desired rate, these results would be statistically significant with a sample size of 12 or more. Similarly, if a researcher planning an audit has set 90% as the acceptable target but expects to observe a compliance rate of only 50%, a sample size of only 6 may be sufficient to confirm that "hunch."

On the other hand, a more precise estimate of the degree of overuse (i.e., showing smaller differences) will require a much larger sample. Many resources and online calculators are available for calculating sample sizes.[9]

Table 1.2
Minimum samples sizes required for improvement projects based on observed and desired system performance

Observed system performance (%)	Desired system performance	
	80%	90%
95	26	140
90	70	Not applicable
85	260	180
80	Not applicable	50
75	280	28
70	80	20
66	45	15
60	25	10
50	12	6
40	10	5
20	5	5

The table shows the approximate sample size required to reject the null hypothesis that observed performance (from an audited sample) is consistent with the desired system performance, shown here as being either 80% or 90%. The results shown here all use the conventional two-tailed p value of 0.05.

Reproduced with permission from Etchells E, M Ho and KG Shojania. 2016. Value of small sample sizes in rapid-cycle quality improvement projects. *BMJ Quality & Safety*. **25**(3):202–6.

Case example 1.1

The use of proton pump inhibitors

A physician working in a family health team recently had a patient admitted to hospital for acute interstitial nephritis related to proton pump inhibitor (PPI) use. She decides to perform a chart review of consecutive patients managed by the team to determine whether they had met at least one of the accepted clinical indications for ongoing PPI use. She decides that if fewer than 80% of patients had a clear indication for PPI use, she will invest time in developing a formal de-prescribing intervention with her colleagues in the practice. After reviewing the first 10 charts, she finds that PPIs were indicated in only 4 (40%) of patients. Despite the small sample, the audit is already sufficient to reject the null hypothesis that performance is at least 80%. Assuming that the audit is taken from a representative sample of patients, it can already be concluded that the performance is substantially below 80%.

Sampling method and data collection

Although a small sample may be all that is needed to confirm resource overuse in some situations, small samples have limitations and it will be important to ensure the internal validity of practice audits.[8] Patient eligibility should be well defined with pre-established exclusion criteria, and a log should be kept to ensure that data collection is complete. The impact of incomplete data on audit results is amplified when the sample size is small. External validity is also important: an unbiased sampling method should be used to ensure that the audit is representative of the population of interest. A consecutive or random sample of patients should be preferred over a convenience sample, since the latter can introduce bias. Etchells and colleagues[8] provide details on the methodology of sampling, as does the Royal College *Research Guide*.[4]

Determining appropriateness criteria

Before an audit is performed, appropriateness criteria for the test, treatment or procedure of interest should be determined. These criteria, which should remain unchanged throughout the study, might be taken from guidelines or the best available research evidence. Using objective criteria (such as laboratory values, medication route or validated risk scores) will facilitate the audit. It is also helpful to share the list of criteria with colleagues from the practice setting beforehand to confirm their agreement and ensure that important indications for the test, treatment or procedure have not been overlooked. This will decrease the chance that the results of an audit overestimate the degree of resource overuse.

Performing the audit

A simple chart abstraction tool (similar to a case report form) is used in most cases to document appropriateness criteria and any other variables of interest. The audit may be retrospective or prospective. The type of data collection may depend on the sample size and feasibility of doing prospective audits, and on whether the clinical documentation reviewed retrospectively can determine, with accuracy, whether patients had an appropriate reason to be tested. Prospective and retrospective audits are described in Case Examples 1.2 and 1.3, respectively.

Case example 1.2

Urinalysis orders in the emergency department[10]

Dr. Penny Yin, a resident in Internal Medicine, notices that urinalysis is ordered for most patients admitted to the General Medicine (GM) service in her hospital and that the results are often abnormal. Because positive results are frequently nonspecific, she wonders whether urinalysis is being overused and is leading to the overprescription of antibiotics. She approaches her supervisor about doing a project to assess the appropriateness of urinalysis orders for GM patients. After reviewing the literature, she develops a list of appropriate indications for urinalysis (e.g., acute kidney injury, rhabdomyolysis, nephrotic syndrome, vasculitis or urinary tract infection [UTI]). Because the documentation of urinary symptoms is known to be poor, she decides to perform a one-month prospective audit of all patients for whom urinalysis is ordered to determine whether they have acute urinary symptoms or any other documented reason for testing. Of 403 consecutive patients admitted to GM (median age, 79 years; 212 [52.6%] women), 250 (62.0% [95% CI 57.3%–66.7%]) had a urinalysis on admission. Of these patients, 211 (84.4% [79.9%–88.9%]) lacked symptoms of UTI, and 198 (79.2% [74.2%–84.2%]) lacked symptoms of UTI and had no other appropriate clinical indication for testing.

Case example 1.3

Routine postoperative radiography after orthopedic trauma surgery

Dr. Michel Taylor, an orthopedics resident, is interested in the value of routinely ordering a postoperative x-ray after orthopedic trauma surgery. He reviews all consecutive patients admitted with a fracture diagnosis who underwent surgical fixation from 12 June 2012 to 30 October 2014. The presence of an immediate complication is determined through the assessment of saved fluoroscopy images and postoperative check x-rays by both a senior orthopaedic resident and a staff radiologist. Discharge notes and postoperative ambulatory clinic notes are also reviewed for each patient. He identifies only two instances (2/1164; 0.1%) in which complications were noted on postoperative x-ray and led to a change in management and a return to the operating room. This audit suggests that routinely ordering x-rays after trauma surgery has a very low yield, and that ordering x-rays on a case-by-case basis might be a higher-value approach.

Module Two: Projects to develop measures for resource utilization

Literature search

As with all scholarly projects, it is helpful to begin a resource stewardship study with a literature search. This can be undertaken in a structured or unstructured way, depending on whether this is the main focus of the project. A review of the literature can uncover what is already known about the utilization of the test, treatment or procedure of interest. The literature may also contain clinical guidelines that define appropriateness criteria for its use. In other cases, the primary literature will need to be reviewed to deduce appropriateness criteria. If observational or intervention studies have already have been performed, it can be helpful to review the methods they used to measure resource utilization (see Case Example 2.1).

Case example 2.1

Appropriateness of magnetic resonance imaging

Dr. Neil Kalra, a radiology resident, has noticed that the wait list for magnetic resonance imaging (MRI) at his hospital is extremely long, and yet some of the MRIs were done without an appropriate clinical indication. He sets out to determine what is known about the appropriateness of MRI testing in Canada. He quickly realizes that little is known about this topic and decides to undertake a systematic literature search identifying studies related to MRI appropriateness in Canada published between 2003 and 2013. He recruits a colleague to help; after establishing inclusion and exclusion criteria, he and his collaborator independently review the literature. Using a predefined search strategy, they arrive at a final sample that contains 8 quantitative studies that report rates of inappropriate MRI exams ranging from 2% to 28.5%. Their review also reveals substantial variations among studies with respect to methods and analyses, including definitions of appropriateness.

Evaluating resource utilization measures

Although the literature may point to standard definitions for and ways of measuring resource utilization, putting any measurement tool into practice in a specific patient population generally requires some evaluation to ensure that it is valid (meaning that it accurately reflects resource utilization) and reliable (meaning that it measures resource utilization in a consistent way). The Royal College *Research Guide*[4] provides more information about the difference between validity and reliability (see chapter 10).

For some projects, the method of measuring resource utilization is straightforward and unlikely to have significant problems with respect to validity or reliability. For example, a project that aims to reduce daily bloodwork may involve a laboratory report that lists the patients who received bloodwork from the patient unit under evaluation. Assuming these reports reflect the patient population of interest, this method of tracking resource utilization is likely to be highly valid and reliable. On the other hand, a project that aim to reduce antibiotic prescriptions for patients who have acute aspiration episodes will require adjudication of appropriateness of antibiotics and a valid and reliable definition for patients with aspiration events.

To test its validity, a measure should be compared with a known gold standard (see Case Example 2.2). To test the reliability of a measure, residents undertaking a resource stewardship project should be encouraged to have two independent reviewers perform an audit of the same patient group, to help determine whether criteria for the appropriate use of a test, treatment or procedure are applied consistently. This exercise allows the calculation of the inter-rater reliability (kappa score) – that is, the level of agreement between reviewers. A high kappa score suggests that the tool is reliable in the patient population under evaluation.

Case example 2.2

Measuring urinary catheter use on inpatient wards[11]

Urinary catheters are frequently inserted and left in place for longer than required among patients in hospital. An improvement team at a teaching hospital has already documented that 18% of their patients are catheterized, and that in over 50% of these patients there is no appropriate clinical indication for doing so. As they begin to consider how improvements might be made, the team realizes that they have no way to measure urinary catheter use over time without performing bedside audits, which are labour intensive and difficult to sustain. In reviewing the literature, they identify studies in which an electronic medical record was used successfully to automate the measurement of urinary catheter use. Since their hospital's electronic database of nursing activities includes the insertion and maintenance of urinary catheters, they decide to test whether these hospital data will be accurate enough (valid) and reproducible (reliable) to measure urinary catheter use. For 12 weeks, they perform random bedside audits to confirm the presence or absence of urinary catheters (the gold standard) and compare their results with the information recorded in the hospital database on the same day. Among 1516 patients, 236 (15.6%) had a urinary catheter on bedside audit; of these, 24 (1.6%) were missed by the electronic system. In addition, 10 "false positives" were found in the databases; these resulted from the system's inclusion of other urinary devices in the count. The improvement team calculates that the sensitivity, specificity and kappa statistic for the measurements derived for the database are 89.8, 99.3 and 0.92, respectively. From these results, the team concludes that the hospital database is a valid way of measuring urinary catheter use over time.

Module Three: Studies to understand patterns or key drivers of overuse

A range of diagnostic tools used in QI methodology can be used to identify and understand the key drivers of resource overuse. Faculty should familiarize themselves with the most common of these tools so they can teach trainees how to apply them in their resource stewardship projects. Depending on the project, one or all of the tools described below may be helpful in determining the main reasons for the wasteful use of tests, treatments or procedures.

Fishbone diagrams

Fishbone diagrams are a simple yet powerful tool for identifying variables that might influence a resource utilization problem. They are generated through a multidisciplinary brainstorming session involving all participants or stakeholders in the process leading to the overuse of a test, treatment or procedure (see Case Example 3.1). The goal is to generate a more comprehensive understanding of the problem than would be achieved by just one or two physicians working alone. In some situations, engaging patients and families in the process can be helpful. Generally, the first step of the brainstorming is to establish broad categories of contributing factors: these form the "bones" of the "fish." Patient, provider, equipment and organizational factors are commonly used categories, but other groupings may apply, depending on the problem.

Case example 3.1

Unnecessary free thyroid hormone testing in patients with normal thyroid-stimulating hormone level[12]

A group of Endocrinology subspecialty residents, including Dr. Julie Gilmour, focused their resource stewardship project on trying to understand the reasons for unnecessary free thyroid hormone testing at their hospital. To identify some of the key factors that might be implicated, the project team held a brainstorming session to create a fishbone diagram. The resulting diagram (Fig. 1) captured a wide range of potential contributors to the problem of unnecessary free thyroid hormone testing. Although the team realized that the relative importance of some of these factors might vary, the purpose of the exercise was to generate a comprehensive list that could then be assessed further using other analytical tools, such as a Pareto diagram (see Fig. 4).

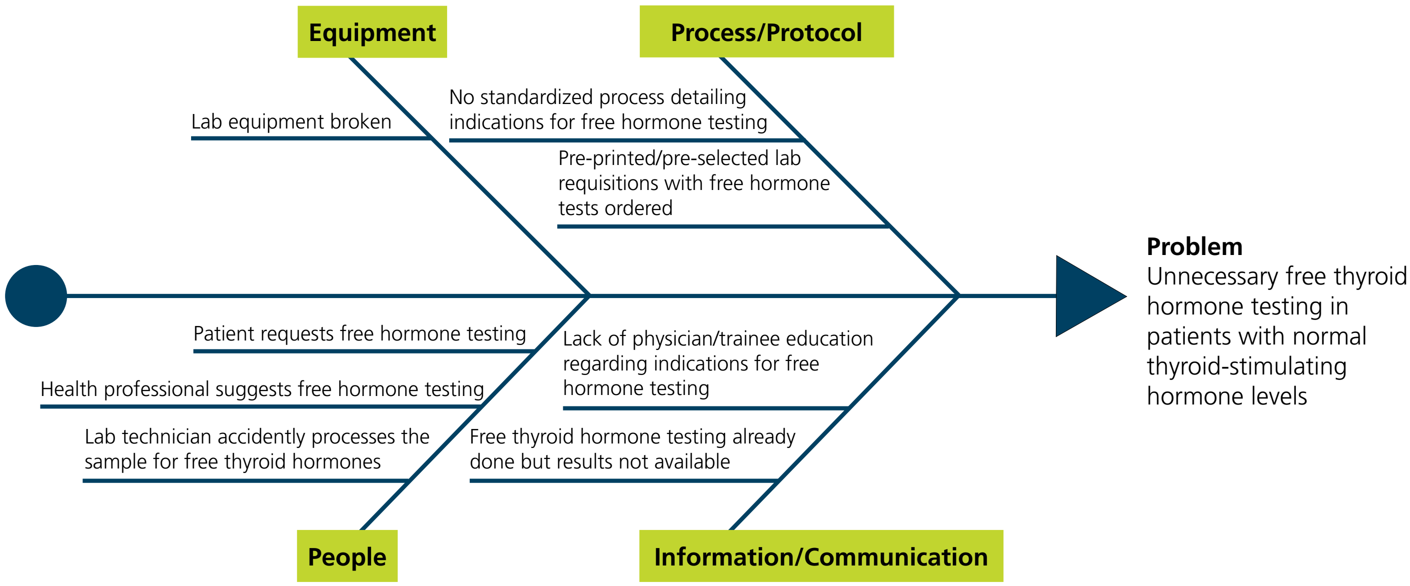


Figure 1
A fishbone diagram

Process mapping

Process mapping is a simple tool that involves creating a diagram of the steps involved in the process that leads to the test, treatment or procedure of interest. Frequently, a process is more convoluted than it might first appear, and so drawing a process map can uncover important steps that generate waste within the system (see Case Example 3.2). The best way to draw an accurate process map is to visit the care setting where the process takes place and follow it, step by step, from start to finish. Moving physically through the process in this way also gives the study team the opportunity to speak to participants and stakeholders involved in the process, and thus understand what is happening to drive overuse.

Case example 3.2

Unnecessary free thyroid hormone testing – continued[12]

Continuing their investigation of unnecessary free thyroid hormone testing, the resident team discovered that free thyroid hormone testing was often being ordered for patients with a normal TSH level, and that in most circumstances the entire thyroid hormone panel was normal (Fig. 2). More specifically, it was noted that 65% of free T_4 (fT_4) and 59% of free T_3 (fT_3) measurements were ordered for patients with a normal TSH level, and that these accounted for a high proportion of “unnecessary” free thyroid hormone tests. A TSH-centred approach to ordering thyroid function tests is advocated by most professional associations; therefore, the team redesigned the laboratory process to provide a fT_4 result only if the TSH was abnormal (Fig. 3). In the new system, a fT_4 or fT_3 test would also be done if clinical justification was provided on the laboratory requisition.

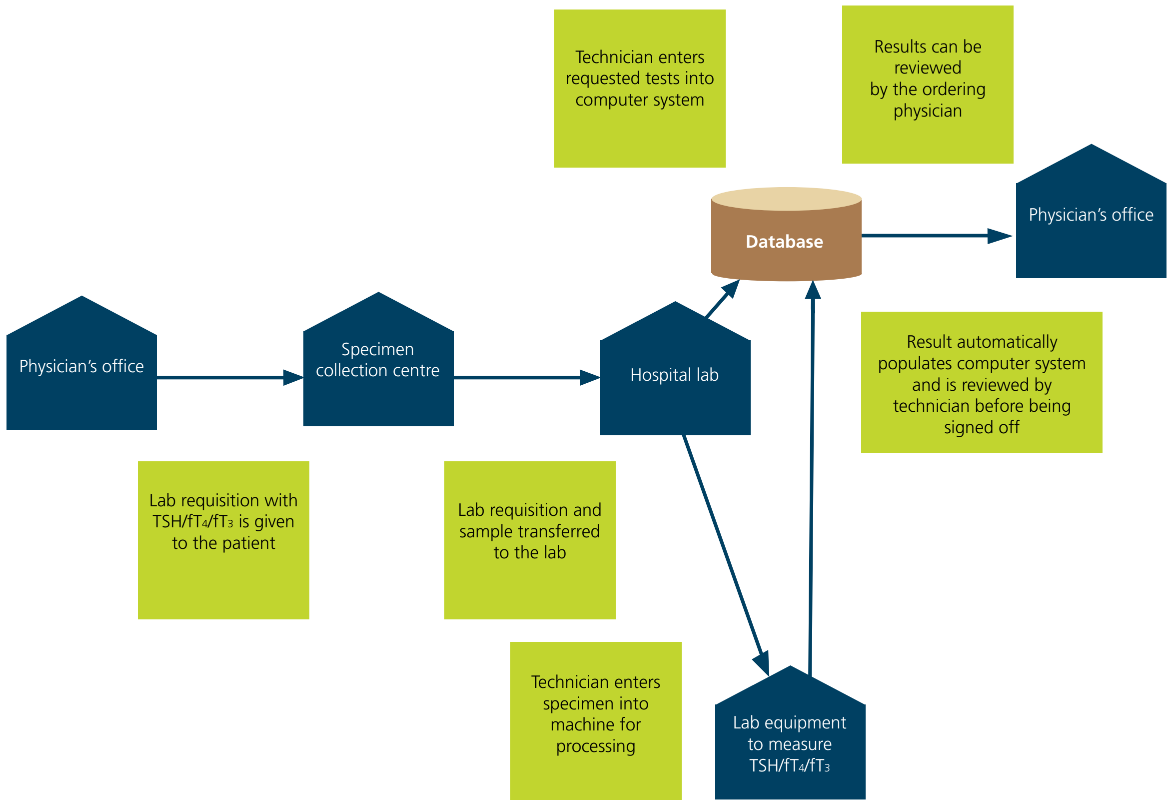
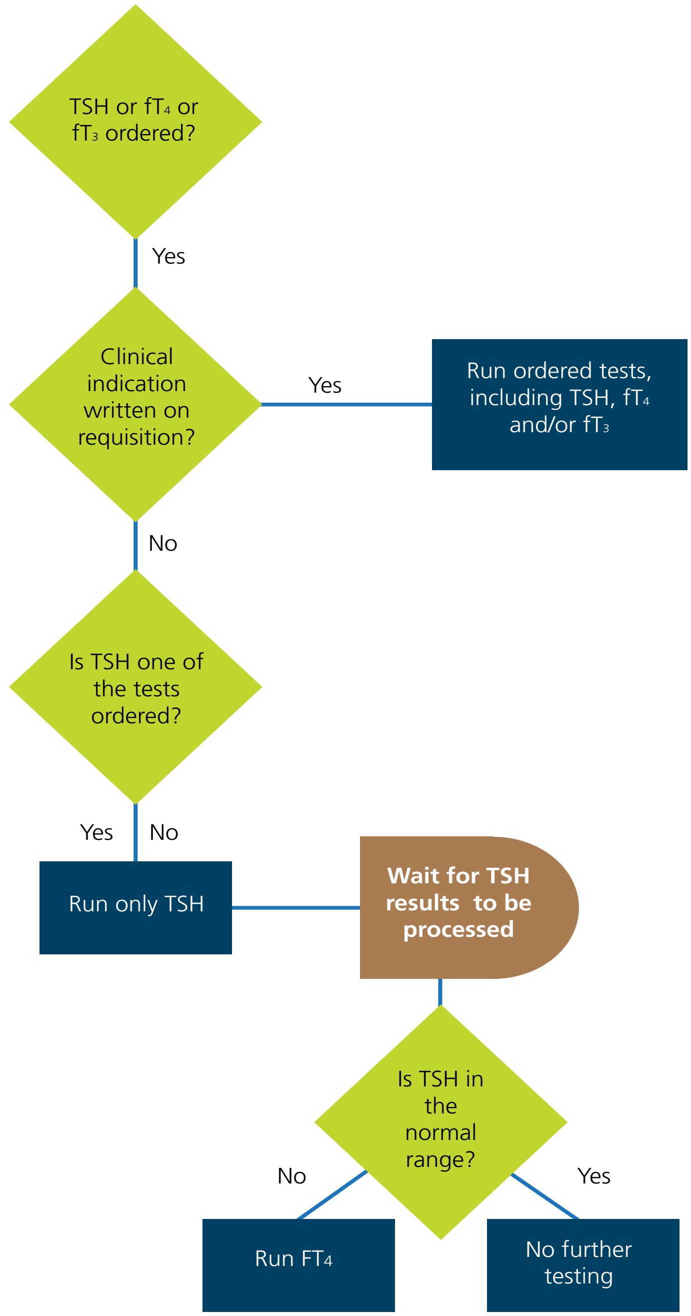


Figure 2

A process map: analysis of steps leading to unnecessary thyroid-stimulating hormone testing



Clinical indications for free hormone testing:

Any of the following:

- Graves' disease
- TSH-producing pituitary tumor
- Liothyronine and/or dessicated thyroid hormone therapy
- TSH resistance syndrome
- Pregnancy

or

- Any manual notations in the clinical indication box by the requesting physician for free thyroid hormone measurement

Figure 3

A process map: thyroid-stimulating hormone and free-thyroid hormone testing after study intervention

Pareto diagrams

A Pareto diagram (Fig. 4) is a type of bar graph that displays values in descending order to determine the cumulative frequency of the most important factors associated with a problem. The idea is to identify the small number of factors that contribute to the largest use of resources. To create a Pareto diagram, at least a dozen variables would be abstracted from the potentially avoidable test, treatment or procedure to see whether there is a strong association between one or two of these variables and resource overuse.

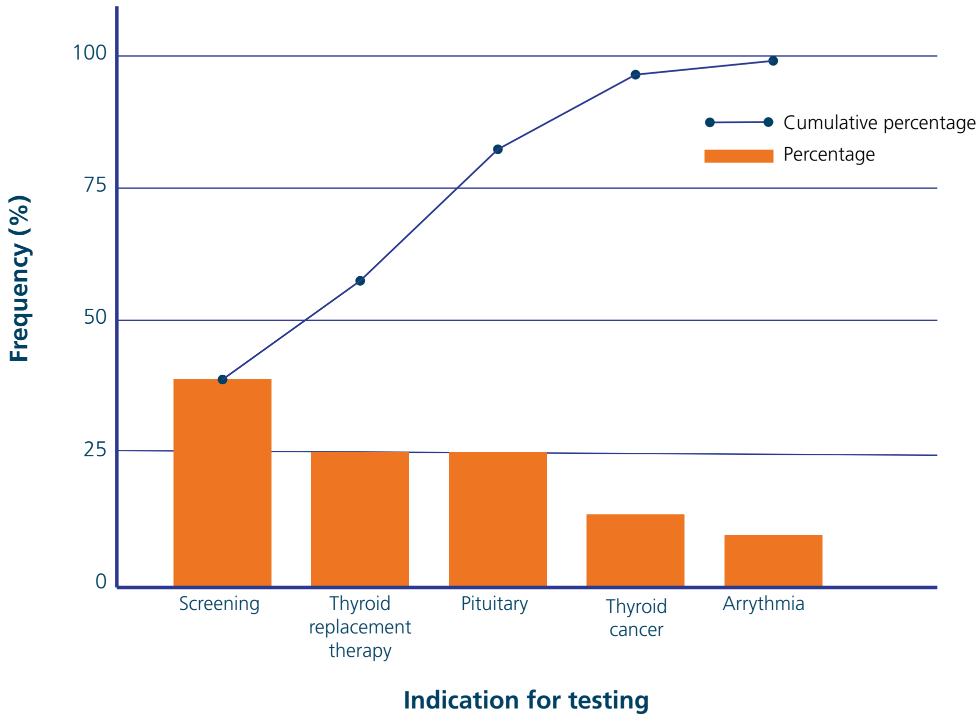


Figure 4
A Pareto diagram

Module Four: Studies undertaken to test a change to reduce overuse

Stakeholder identification and engagement

Resource stewardship projects are often challenging, as they cannot be performed without the active engagement of key stakeholders – the individuals who will be affected by the changes being considered. Remind your residents that the crucial first steps in this type of project are to identify all stakeholders and to assemble a project team that will ensure strong representation of these stakeholders and facilitate the open sharing of ideas. Although trainees may lead the project, the engagement of stakeholders will often rely on faculty, who are in a better position to ensure that representatives from key stakeholder groups are present and willing to participate.

Once stakeholders are brought together, getting them to work through the diagnostic tools described in module 3 – e.g., brainstorming a fishbone diagram – is a great way to engage everyone in the process and to gain insight from their perspectives into factors that might otherwise be overlooked.

Articulating a hypothesis

Making a change to mitigate resource overuse requires an understanding of the key system factors that promote the excessive use of tests and procedures. To increase the chance of success, the changes made need to be well matched to the problem identified. The diagnostic tools described in module 3 can help identity the problem and support the articulation of a hypothesis linking it to the change idea. For instance, if the main driver of inappropriate orders for partial thromboplastin time (PTT) testing is the fact that this test is bundled in the order-entry system with orders for prothrombin time (PT) tests, such that every time a physician orders a PT test he or she also receives a PTT test, then a clearly linked change idea would be to unbundle these tests within the order-entry system. Case Example 4.1 illustrates the importance of importance of a well-founded hypothesis.

Case example 4.1

Avoiding unnecessary urinary catheter insertion in the emergency department

A hospital develops an educational campaign to reduce the use of urinary catheters in its emergency department. Three months into the campaign, audits reveal that urinary catheter insertions are still occurring among patients who lack an appropriate clinical indication. A focus group discussion reveals a lack of urinals and commodes in the emergency department, along with a shortage of support workers to assist with alternative methods of toileting. This example highlights the need to articulate a hypothesis that links the intervention to the change idea before investing efforts in implementation. The use of an educational intervention would be appropriate only if a knowledge gap was the underlying problem. In this case, equipment factors were likely the main contributor.

Test of change

After a hypothesis has been articulated, a test of change can be made to determine whether the intervention has the anticipated effect. In the most common model of improvement, known as the “plan–do–study–act” (PDSA) cycle, changes are made with iterative modifications until the desired outcome is achieved. This method begins with a prediction (plan), tests the change at a small scale (do), reviews the results (study), and then determines what modifications need to be made to the hypothesis and the next test of change (act). Because changes rarely result in improvement the first time, multiple PDSA cycles are usually required. For a practical “primer” on implementing a PDSA cycle, see Leis and Shojania 2017.[13]

Measuring change

Each PDSA cycle may use different types of data to explore the problem and test the impact of various change ideas. As the project evolves, the following types of longer-term measures typically used in health care improvement studies should also be applied:

- **Outcome measures:** to determine the effect of the change on rates of use of the test, treatment or procedure under study (e.g., reduction in rates of inappropriate urinary catheter insertions)
- **Process measures:** to determine whether the steps in the new process are being followed as intended (e.g., number of physicians correctly completing a risk-scoring tool for medical inpatient venous thromboembolism prophylaxis)
- **Balancing measures:** to determine whether changes made in one part of a system are resulting in unintended consequences in another part of the system (e.g., an intervention to eliminate daily blood work might assess for any increase in stat complete blood counts, as a way of looking for evidence that patients who needed this test were being missed as a result of the intervention).

Resource stewardship projects differ from other health care projects in that they are focused on reducing testing, treatment and procedures as their main outcome measure. Whether this reduction represents an improvement to the system overall cannot be assessed without an accompanying balancing measure (see Case Example 4.2). In other words, decreasing resource utilization alone is not adequate, as particular care must be taken to look for any potential patient harm related to the decrease in investigations and treatment.

Case example 4.2

Elimination of pre-operative screening urine cultures[14]

Screening patients who undergo non-urologic surgery for the presence of asymptomatic bacteriuria is a controversial practice, as evidence linking this treatment to the prevention of surgical-site infections is lacking. One hospital implemented a policy to stop processing screening urine culture specimens from a pre-operative clinic, except in the case of telephone requests made for patients with a symptomatic urinary tract infection. As his research elective, Dr. Michael Lamb, a resident in Internal Medicine, led a study to evaluate the impact of this new system of care. The primary outcome measure was the number of urine cultures ordered, and the process measure was the number of urine cultures processed (the elimination of almost all of these would indicate that the laboratory was following its new policy). As balancing measures, a telephone log was kept to track the number of phone calls received from the pre-operative clinic, and prospective surveillance for surgical-site infections was performed to determine whether any increase in these infections occurred after the intervention. The change was associated with a 99% reduction in urine cultures processed, a single call to the laboratory for a patient with UTI, and no significant increase in the incidence of prosthetic joint infection.

Module Five: Projects to evaluate the safety and sustainability of a new process of care

Once a new process to reduce resource waste has been tested, the next step is to determine whether it can be safely and sustainably implemented at a larger scale and/or over a longer period. Because of the resources involved in conducting these studies, larger-scale evaluations of a new process of care should not be undertaken until the intervention has been fully developed and successfully implemented in a pilot study.

Evaluative design

Resource stewardship interventions can be evaluated as robustly as any other QI intervention. Various randomized and non-randomized designs are possible, although non-randomized study designs are more commonly used in this type of research. Below is a brief description of these evaluative methods, which are reviewed in more detail elsewhere.[2,15] A helpful overview of research designs used to evaluate health system interventions, including their presentation of data, is provided by Coly and Parry 2017.[16]

- **Stepped-wedge design.** A resource stewardship project using this study design would implement the sequential rollout of an intervention among clinicians, units or institutions over a number of periods, so that by the end of the study all participants have been exposed to the intervention. This design gets its name from the interlocking wedge shapes created by diagrammatic representations of the rollout. The order in which participants receive the intervention may be randomized. Data are collected and outcomes measured at each point at which a new group of participants receives the intervention. Observed differences in outcomes between the control (pre-intervention) section of the wedge and those in the intervention section are assumed to be attributable to the intervention.

- **Time-series design.** In resource stewardship projects using this type of design, data are collected and outcomes measured at multiple points before and after the introduction of the intervention. Data collected at multiple points before the intervention allow the underlying trend and any secular effects to be estimated. Data from multiple points after the intervention allow the intervention effect to be estimated while accounting for the underlying secular trends. The intervention may be stopped and re-started a number of times, or the intervention may be implemented just once.
- **Controlled before-and-after studies.** In this study design, a control group with characteristics and performance similar to those of the group receiving the intervention is included. Data are collected and outcomes measured in both groups before and after the intervention is introduced in the study group. Observed differences between groups in the intervention period (from baseline in each group) are assumed to be attributable to the intervention.
- **Uncontrolled before-and-after studies.** In before-and-after studies conducted without a control group, outcomes are measured in a study setting before and after an intervention is introduced. In this case, the period before the intervention acts as the "control." Observed differences in the outcomes are assumed to be attributable to the intervention.

Measuring data over time

Displaying data over time provides for a much richer analysis than simply presenting before-and-after comparisons of aggregate results, as might be shown on a histogram. As in other QI projects, run charts and statistical process control charts (SPC) can be used to display data over time. Table 5.1 summarizes the main differences between run charts and SPC charts.

Table 5.1
Comparison of run charts and statistical control process (SPC) charts

	Run chart	SPC chart
When to use	For early signal recognition (e.g., 8–12 data points)	To study changes over time (e.g., minimum of 15–20 data points)
How to apply	Any data can be used to create a run chart; data are plotted over time and a median line is determined.	The type of SPC chart used depends on the type of data examined. A mean is established and drawn as a central line; upper and lower control limits are then set as a fixed number of standard deviations from the mean.
How to interpret	Run chart rules are applied to the data.[17]	The presence of special-cause variation is determined by available software.

Run charts can be created using any type of data measured over time, and the application of "run chart rules" are a simple way of determining whether statistically significant changes have occurred. Perla and colleagues[17] provide a full description of how to apply these rules to interpret run charts.

SPC charts are more specific to the type of data being collected and require a larger sample (at least 15–20 data points over time). A published example of a P-chart (a type of SPC chart) is shown in Figure 5. It represents the proportion of patients who underwent elective joint arthroplasty who had a pre-operative screening urine culture from May 2013 to June 2016.[14] The control limits (dashed lines) were set at 3 σ (equivalent to 3 standard deviations) from the mean that was determined in the pre-intervention phase (May 2013–June 2015), such that consecutive points below the lower control limits represent a statistically significant decrease. The figure shows a 2-month washout period (May–June 2015). An abrupt decease in urine culture orders is seen starting in July 2015, when the intervention was initiated. The control limits continue at the pre-intervention baseline. From the consecutive points below the lower control limits from the beginning of the intervention, we can conclude there was a statistically significant decrease in urine cultures ($p < 0.01$). These methods are described in more detail elsewhere.[18]

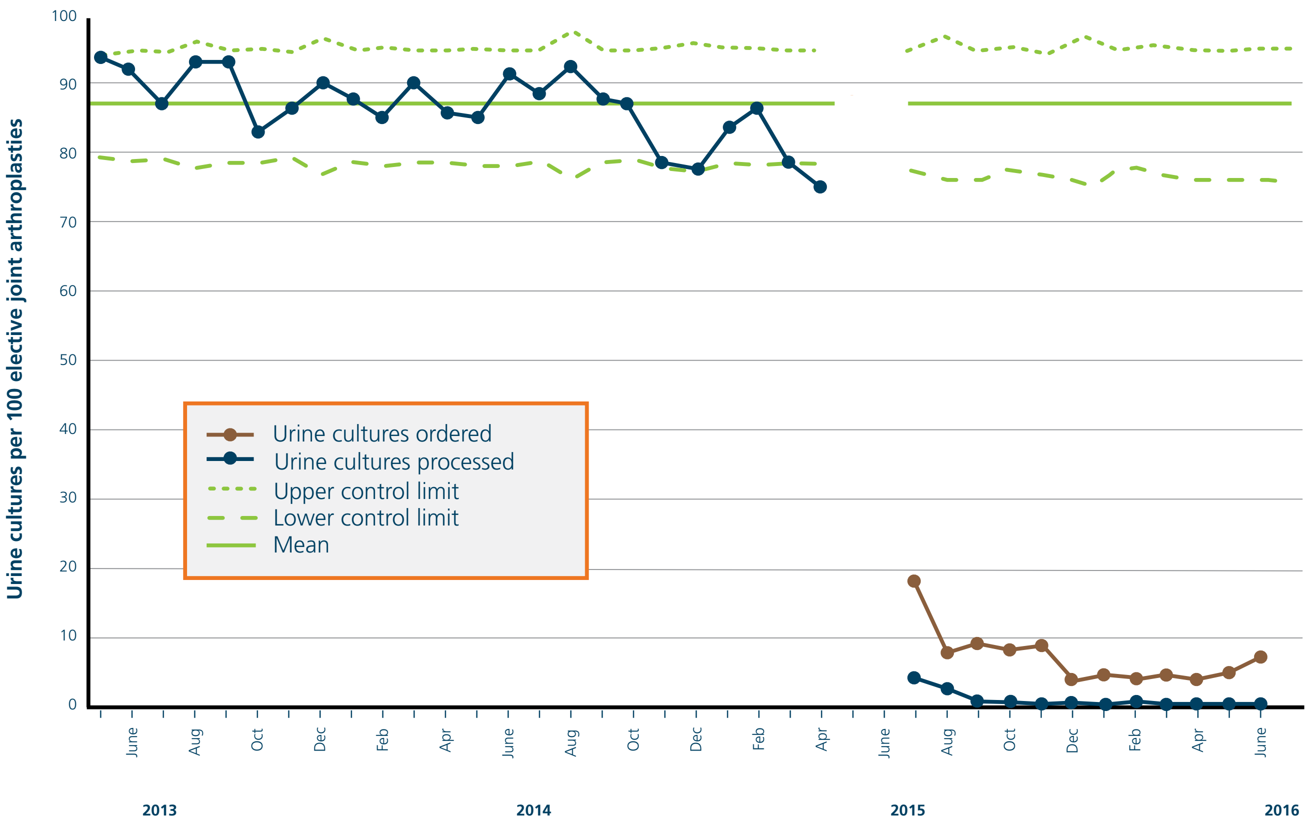


Figure 5

Example of a process control chart. The figure shows the monthly average number of urine cultures ordered and processed per 100 elective arthroplasties before and after implementation of a policy to no longer process screening using cultures. Adapted with permission from Lamb et al. 2017.[14]

ASSESSMENT TOOL FOR RESOURCE STEWARDSHIP PROJECTS

To reflect the specific skill set demanded by resource stewardship, these projects should be assessed and evaluated differently from other research and health care projects. The following assessment tool can be used to assess competencies related to the ability to undertake resource stewardship projects.

ASSESSMENT FORM: RESOURCE STEWARDSHIP PROJECT

Title of project:

Team members:

Rating system

0 = No
1 = Some attempt was made but does not meet the requirements.
2 = Met some requirements but substantial improvement is required.
3 = Good (can use some improvement).
4 = Very good (only minimal improvement is required).
5 = Excellent (no improvement needed).
NA = Not applicable (out of scope for resource stewardship project).

Please circle appropriate number for each question							
1. Is the scope of the resource stewardship project clearly defined?	0	1	2	3	4	5	NA
2. Have all key stakeholders been identified and involved wherever appropriate in the project design and implementation?	0	1	2	3	4	5	NA
3. Has the need for research ethics board approval been clearly addressed?	0	1	2	3	4	5	NA
4. Are appropriateness criteria clearly defined?	0	1	2	3	4	5	NA
5. Does the project include an appropriate family of measures (including one or more balancing measures)?	0	1	2	3	4	5	NA
6. Are there prespecified exclusion criteria and is the data collection complete?	0	1	2	3	4	5	NA
7. Has a hypothesis about the key drivers of overuse been articulated?	0	1	2	3	4	5	NA
8. Have tests of change been performed with demonstrated modifications to the intervention or implementation strategy (i.e., demonstration that PDSA was executed)?	0	1	2	3	4	5	NA
9. Has the project led to sustainable system change to promote better resource utilization?	0	1	2	3	4	5	NA
10. Has any potential harm associated with reduced utilization been identified and mitigated where appropriate?	0	1	2	3	4	5	NA

Comments

Score ___ **/50**

Adapted from Wong RYM. 2015. *Teaching quality improvement in residency education*. Ottawa: The Royal College of Physicians and Surgeons of Canada. p 102.

TIP SHEET FOR SUPERVISING A SUCCESSFUL RESOURCE STEWARDSHIP PROJECT

✓ **Select a relevant topic.** Sometimes tests or treatments are perceived to be overused when in fact any overuse is minimal or has little impact. A review of the literature and/or local data can help confirm whether a perceived issue does in fact present a worthwhile opportunity for resource stewardship research and intervention.

✓ **Create an improvement team.** Resource stewardship projects are almost never undertaken in isolation. Although a trainee may be given the lead, the role of faculty is to ensure that key stakeholders become involved early in the project. Sometimes interdisciplinary supervision can help ensure that there is a more complete understanding of the problem and that all stakeholders help to design the changes proposed.

✓ **Avoid scope creep.** Resource stewardship projects have a tendency to expand, involving more tests and broader patient populations than in their original conception. Clearly defining the scope of the project with regard to the test, treatment or procedure of concern, the patient population involved, and the specific type of project that is being undertaken (e.g., defining the problem) will help to ensure that the project is completed within the allotted time.

✓ **Build on existing projects.** Taking a resource stewardship project from start to finish is not always realistic. It may be feasible for one resident to conduct the first phases of a project and then hand it on to other trainees for completion. For example, a project that has revealed the key drivers of overuse and articulated a strong hypothesis could be handed off to another trainee to develop tests of change at a small scale, that could in turn could provide the basis for a pilot project to be conducted by others.

✓ **Block off time to meet with trainees regularly.** Taking on a resource stewardship project can be daunting, as it often means questioning the status quo and thinking innovatively about solutions. Trainees will encounter roadblocks along the way that will require faculty input as to how best to proceed. Scheduling regular meetings with trainees to monitor progress will help maintain momentum and ensure that educational needs are met.

✓ **Have a protocol.** As in other scholarly projects, a protocol for a resource stewardship project should be developed. This will serve as the roadmap for the project but will also set clear expectations for the work to be completed. It will also identify areas where additional supports may need to be provided, such as access to administrative data.

REFERENCES

1. Frank JR, L Snell L and J Sherbino, editors. 2015. *CanMEDS 2015 Physician Competency Framework*. Ottawa: The Royal College of Physicians and Surgeons of Canada.

2. Health Quality Ontario. 2012. *Quality Improvement Guide*. Toronto: Queen's Printer for Ontario.

3. Wong RYM. 2015. *Teaching quality improvement in residency education*. Ottawa: The Royal College of Physicians and Surgeons of Canada.

4. Harvey BJ, Lang ES and Frank JR, editors. 2011. *The research guide: a primer for residents, other health care trainees, and practitioners*. Ottawa: The Royal College of Physicians and Surgeons of Canada.

5. World Health Organization. 2006. *Quality of care: a process for making strategic choices in health systems*. Geneva: WHO Press. Available from: http://apps.who.int/iris/handle/10665/43470

6. Alberta Research Ethics Community Consensus Initiative (ARECCI) Network. 2005, revised 2010. ARECCI Ethics Screening Tool. Edmonton: Alberta Innovates. Available from: www.aihealthsolutions.ca/initiatives-partnerships/arecci-a-project-ethics-community-consensus-initiative/tools-and-resources/

7. Choosing Wisely Canada. *Recommendations and Resources, by Specialty*. Available from: www.choosingwiselycanada.org/recommendations/

8. Etchells E, M Ho and KG Shojania. 2016. Value of small sample sizes in rapid-cycle quality improvement projects. *BMJ Quality & Safety*. **25**(3):202–6.

9. Brant R. *Sample size calculators*. Available from: www.stat.ubc.ca/~rollin/stats/ssize/index.html

10. Yin P, A Kiss and JA Leis. 2015. Urinalysis orders among patients admitted to the general medicine service. *JAMA Internal Medicine*. **175**(10):1711–3.

11. Leis JA, C Corpus, B Catt, F Jinnah, B Edgar, BM Wong, S Callery, AE Simor and M Vearncombe. 2015. Indwelling urinary catheter surveillance using a task-oriented nurse acuity system. *American Journal of Infection Control*. **43**(10):1112–3.

12. Gilmour JA, A Weisman, S Orlov, RJ Goldberg, A Goldberg, H Baranek and G Mukerji G. 2017. Promoting resource stewardship: reducing inappropriate free thyroid hormone testing. *Journal of Evaluation in Clinical Practice*. **23**(3):670–5.

13. Leis J and KG Shojania. 2017. A primer on PDSA: executing plan-do-study-act cycles in practice, not just in name. *BMJ Quality and Safety*. **26**(7):572–7.

14. Lamb MJ, L Baillie, D Pajak, J Flynn, V Bansa, A Simor, M Vearncombe, SA Walker, S Clark, J Gollish and JA Leis. 2017 Elimination of screening urine cultures prior to elective joint arthroplasty. *Clinical Infectious Diseases*. **64**(6):806–9.

15. Fan E, A Laupacis, PJ Pronovost, GH Guyatt and DM Needham. 2010. How to use an article about quality improvement. *JAMA*. **304**(20):2279–87.

16. Coly A and G Parry. 2017. *Evaluating complex health interventions: a guide to rigorous research designs*. Washington (DC): AcademyHealth. Available from: www.academyhealth.org/files/AH_Evaluation_Guide_FINAL.pdf

17. Perla RJ, LP Provost and SK Murray. 2011. The run chart: a simple analytical tool for learning from variation in healthcare processes. *BMJ Quality and Safety*. **20**(1):46–51.

18. Provost LP and S Murray. 2011. *The health care data guide: learning from data for improvement*. San Francisco: Jossey-Bass.

CONTRIBUTORS AND REVIEWERS

Contributors

Julie Gilmour, MBChB, MSc, HQ, FRCPC
Department of Endocrinology
University of Toronto
Toronto, Ontario

Heather Johnson, MD
Department of Internal Medicine
Queen's University
Kingston, Ontario

Neil Kalra, MD
Department of Diagnostic Radiology
University of Saskatchewan
Saskatoon, Saskatchewan

Michael Lamb, MBBS
Department of Internal Medicine
University of Toronto
Toronto, Ontario

Jennifer Ngo, MD, MSc (QI/PS), FRCPC
Department of Medicine
University of Calgary
Calgary, Alberta

Michel A. Taylor, MD, FRCSC
Department of Orthopaedic Surgery
Queen's University
Kingston, Ontario

Mina Tohidi, MD
Department of Orthepedic Surgery
Queen's University
Kingston, Ontario

Penny Yin, MD, FRCPC
Department of General Internal Medicine
University of Toronto
Toronto, Ontario

Reviewers

Stacy Ackroyd-Stolarz, BSc(OT), PhD
Department of Emergency Medicine
Dalhousie University
Halifax, Nova Scotia

Anees Bahji, MD
Department of Psychiatry
Queen's University
Kingston, Ontario

Nancy Fowler, MD, CCFP, FCFP
Academic Family Medicine
The College of Family Physicians of Canada
Ottawa, Ontario

Christopher Moriates, MD
Department of Internal Medicine
The University of Texas at Austin
Austin, Texas

Jennifer Russell, MD, FRCPC
Department of Paediatric Cardiology
University of Toronto
Toronto, Ontario

Martina Trinkaus, MD, FRCPC
Division of Hematology and Medical Oncology
University of Toronto
Toronto, Ontario

Brian M. Wong, MD, FRCPC
Department of Medicine
Sunnybrook Health Sciences Centre
University of Toronto
Toronto, Ontario

Michelle Zeller, MD, FRCPC, MHPE, DRCPSC
Division of Hematology and Thromboembolism
McMaster University
Hamilton, Ontario