Summer 8-8-2020

# A Measure to Determine Acceptable Workload for Increasing Operational Efficiencies for the Conduct of Clinical Trials

Lester Rey Lledo RN, CRNP, DNP
*Duquesne University School of Nursing*

Catherine Johnson FNP, PNP, PhD
*Duquesne University*

Anne Chew PhD
*University of Pennsylvania*

Joan Gilmore BS
*University of Pennsylvania*

Karen Dengel BSN
*University of Pennsylvania*

*See next page for additional authors*

## Author

Lester Rey Lledo RN, CRNP, DNP; Catherine Johnson FNP, PNP, PhD; Anne Chew PhD; Joan Gilmore BS; Karen Dengel BSN; and Jane Anderson MS

# A Measure to Determine Acceptable Workload for Increasing Operational Efficiencies for the Conduct of Clinical Trials

Lester R. Lledo

Duquesne University

**Abstract**

As the demands for the conduct of clinical trials rise, it becomes increasingly important to establish a quantitative means of estimating the appropriate staffing resource to coordinate trial related activities. There has been a limited amount of research conducted to evaluate methods or tools to measure workload in the clinical trial setting. A literature search revealed a gap in the literature about tools used to measure workload and its impact on clinical trial performance and job satisfaction. The aim of this pilot project was to test the Ontario Protocol Assessment Level (OPAL) complexity rating tool in generating quantitative measurements of workload for the purpose of assessing operational efficiencies and identifying opportunities for process improvement changes. The pilot project was conducted in a clinical trials unit consisting of clinical research nurses (CRN), clinical research coordinators (CRC) and research managers who participated in the project implementation. Concepts from the Institute for Healthcare Improvement (IHI) Model for Improvement and Plan-Do-Study-Act (PDSA) were used to guide project implementation and Lean principles for the interpretation of data findings. The findings from the data showed that OPAL can be used as a quantitative means to measure workload, and to assess factors affecting operational efficiencies.

*Keywords: workload, workload measurement, clinical trials, clinical research coordinator, clinical research nurse, skill-mix*

# A Measure to Determine Acceptable Workload for Increasing Operational Efficiencies for the Conduct of Clinical Trials

According to Casner and Gore (2010), "[h]umans who are overburdened with work tend to hurry their performance, commit more errors, yield poor accuracy, become frustrated, uncomfortable, and fatigued, and have poor awareness of their surroundings"…"[i]nterestingly, humans who are underworked can exhibit many of the same symptoms" (p. 1). Therefore, a balanced distribution of work would likely lead to greater productivity and efficiency. They emphasized the need for a device to measure workload that "would allow us to approach any work situation and acquire a numerical (or at least ordinal) measure of the level of workload being experienced by a human operator" (Casner & Gore, 2010, p. 1). In parallel, they stated the need to define "practical and sensible limits for workload", emphasizing that in order to make scientific inferences about numerical measurements of workload levels that these terms need to be more rigorously defined (Casner & Gore, 2010, p. 1).

According to Milani et al. (2017) "few resources are available to quantify clinical trial-associated workload, needed to guide staffing and budgetary planning" (p. 1). They further stated, that "the total number and frequency of procedures specified in each clinical trial protocol…is increasing annually" and affecting the workload of clinical research personnel required to perform these activities (Milani et al., 2017, p. 1). The lack of a method for measuring the needs of patients participating in clinical trials hinders the ability of managers in making precise staffing and budget projections and resource allocation decisions (Brennan et al., 2019). Additionally, Brennan et al. (2019) indicated that when aspects of workload that affect workflow and efficiency are not measured, these have the potential in impacting quality of care, patient safety, and the integrity of research outcomes. This pilot project was implemented in a

clinical trials unit to test and evaluate a workload measurement instrument for generating workload related metrics that could be used to assess factors that impact staff performance and job satisfaction and identify opportunities for process improvement changes to enhance overall clinical trial performance.

## Clinical Research Problem

The basis for this pilot project was the recurring need for research managers of a clinical trials unit to accurately predict staffing requirements and assess work capacity of clinical research staff. As a result, the research staff can more efficiently coordinate complex phase I cancer trials for novel high-risk cell-based therapies (i.e., genetically modified CAR T cells) used to treat a variety of cancer types – hematologic and solid tumor malignancies. This notion provides a more precise estimation of the work capacity of research staff that facilitates a more balanced distribution of work, resulting in greater staff productivity, efficiency and effectiveness, better overall clinical trial performance, and a reduction in work-related stress. Additionally, workload measurements could inform more efficient and cost-effective utilization of staff in accomplishing clinical trial deliverables and meeting project timelines. Historically, research managers from this clinical trials unit have relied primarily on their intuition and experience in assessing feasibility for trial activation and estimating staff capacity for allocation of work assignments. Feasibility assessments are largely based on the projected trial participant sample size and intensity of study visit requirements. An accurate estimation of staffing requirements ensures a safer conduct of clinical trials, high-quality study outcomes, enhances staff efficiency and productivity, and reduces work-related stress and operational costs. For many cancer patients with late stage or advanced disease, participation in clinical trials provide them access to alternative investigational therapies in addition to routine care cancer therapies or in place of

standard of care modalities that are no longer effective in preventing the progression of their disease or resulting in durable remission of their cancer. Administration of investigational therapies can provide benefits in slowing down disease progression, extending quality of life or in some cases resulting in complete remission of cancer.

According to Malik and Lu (2019), "[p]hase I protocols are known to be the most complex and burdensome to conduct" citing that "[c]ontemporary therapies in cancer phase I trials are subject to newer endpoints with an effort to demonstrate a response signal or at least evidence of target inhibition" (p. 519). They evaluated a total of 102 phase I protocols that were active in 1996, 2006 and 2016 and found that there were significantly higher numbers of trial procedures in the protocols from 2016 compared with 2006 and 1996 ($P<0.001$); 90% of these trials were testing immune or targeted therapies (Malik & Lu, 2019). Drug development in oncology continues to grow as the standard treatment for many cancers rapidly evolves to include novel targeted therapies, immunotherapies, antibody-drug conjugates and chimeric antigen T cell receptors (CARs) (Malik & Lu, 2019). The Food and Drug Administration is granting breakthrough designations and conditional approvals based on phase I and II clinical trial data, and as such the development process "should be rapid, efficient and able to implement modern drug development models for upcoming novel drugs" (Malik & Lu, 2019, p. 519). However, "[a]n efficient and successful phase I research program can be challenging in light of an increase protocol complexity, restrictive inclusion and exclusion criteria, high personnel workload, stringent regulatory criteria and a restrictive budget to manage the work" (Malik & Lu, 2019, p. 519).

It is important to quantify workload because the design of phase I trials have substantially increased study requirements making them more complex over the last 20 years (Malik & Lu,

2019). In a study conducted to measure protocol design trends by the Tufts Center for the Study of Drug Development (CSDD), they observed that the number of unique procedures had increased at the annual rate of 6.5% and the frequency of procedures per protocol by 8.7% (Getz et al., 2008). In the same study, they found that "investigative site work burden to administer each protocol increased at an even faster rate of 10.5%" (Getz et al., 2008, p. 450). They described how these protocol changes over time have had an impact on clinical trial performance but have not been sufficiently quantified (Getz et al., 2008). According to Milani et al. (2017), "[t]he assessment of the correct balance between size of workforce and number of trials is therefore essential to ensure on the one hand patient safety and on the other data quality" (p. 7).

## Project Purpose

The purpose of the pilot project was to evaluate the use of a quantitative measurement tool for assessing clinical trial workload, and to identify opportunities for process change and improvement, while promoting job satisfaction.

### *Project AIMS*

1. Utilize OPAL to assess workload of research personnel involved in clinical trial implementation
2. Use data from OPAL to evaluate operational efficiencies impacting clinical trial performance
3. Identify other factors affecting staff efficiency and job satisfaction
4. Use project findings to inform about quality improvement strategies and enhancing job satisfaction

***Project Objectives***

1. Use PDSA cycle to test and study OPAL as a method for measuring clinical trial complexity and workload
2. Use PDSA cycle to assess three factors from OPAL measurements that increase staff effort and impact operational efficiency
3. Use PDSA cycle to
    - Identify and evaluate three dimensions from a job stress survey that impact staff efficiency and contribute to job stress
    - Assess other factors identified from a review of protocol deviations reported for four of the ten protocols in our project sample that may affect staff efficiency and job stress
4. Use Lean principles to interpret project findings and make recommendations for process improvement changes and goals, and improving job satisfaction

## Review of the Literature

A literature search was conducted using PubMed, CINAHL, Google Scholar, Google and Bing. The initial search used terms such as staffing ratio, staffing models, clinical research nurse, clinical trial and caseload. Emulating the concept of the nurse-to-patient ratio, the search focused on methods to determine a balanced distribution of trial protocols assigned to research personnel. The early search generated studies conducted in more traditional and acute patient care settings, such as the intensive care, pediatrics and medical-surgical units. These studies evaluated methods for quantifying nurse staffing requirements based on ratio of patients to nurse or nursing hours per patient day. Additionally, some studies suggested consideration of skill-mix utilizing unlicensed assistive personnel to perform duties or tasks not requiring the completion by a

registered nurse. Although the methods and findings from these studies could not be directly correlated to the circumstances that were specific to the clinical trial setting, the concept of balanced work distribution and skill-level mix was relevant.

Search terms were later refined to include workload, workload measurement, staffing model, staffing ratio, skill-mix and clinical trials. Studies assessing workload and evaluating workload measurement tools specific to the clinical trial setting were identified through a snowball search. A total of six level II and level III good and high-quality studies were identified and determined to be directly related to the project topic and adequately supported the aims for this pilot project. However, it is notable that a gap in the literature for studies evaluating workload measurement tools and the impacts that workload has on clinical trial performance still exist.

A thorough review and appraisal of the literature that directly supported the aims of this pilot project were performed. A summary of the synthesis and appraisal of the literature is provided in Appendix A. The appraisal of the level and quality of the literature was performed according to the criteria described in Johns Hopkins Nursing Evidence-Based Practice: Models and Guidelines (Dang & Dearholt, 2018).

## Synthesis of the Literature

Due to the unique methods described in each study for the measurement of workload, it was important to include a summary of each tool or instrument described to provide better context for the literature synthesis.

In a study conducted by Smuck et al. (2011), they tested the Ontario Protocol Assessment Level (OPAL) complexity rating tool to measure trial complexity, case and total workload. Their study was aligned with the Ontario Institute for Cancer Research's mission for "process of

improvements necessary to promote speed, quality, and accessibility of clinical trials for patients in Ontario" and to assist clinical research professionals in working together to navigate through the very complicated demands for conducting clinical research (Smuck et al., p. 80). OPAL measured trial complexity and assigned a complexity level score based on criteria listed on a pyramid scale. A working group of experienced clinical trial managers was assembled to develop the standard rating scale for their tool to measure trial complexity that would later be applied to measuring workload. The guiding principles for their tool were its ease of use, specificity to measure work directly related to clinical trial professionals, usefulness in calculating workload capacity for different phase trials but not be intended for academic review. The tool is based on a pyramid scale that rates the complexity level of trial protocols from the least to the most complex. They conducted the pilot testing for this tool across several of Ontario's clinical trials sites with the aim of demonstrating that "OPAL was reliable, and that scoring was consistent across clinical trial protocols and across sites" (Smuck et al., 2011, p. 81). Their analysis of the scores from 176 OPAL assessments left them to conclude that a variance of up to 1.5 between the scores was acceptable, with only two trials shown to have a variance of 3.5. The acceptability of a 1.5 variance was based on the research practice models of each site and correlative studies associated with the protocol. Their review of the scores revealed that misinterpretation of the definitions of the parameters used to rate the complexity levels resulted in higher OPAL scores. Hence, they emphasized that "OPAL needs to be applied consistently at the site and should be based on local practices to produce measurable site workload information" (Smuck et al., p. 81). At the conclusion of their pilot project OPAL was revised to more accurately reflect the clinical trial activity reported by the sites. OPAL provides "an objective method of quantifying clinical trials activity on the basis of factors that contribute to increased

workload," however, "total workload of staff varies, and workload needs to be reviewed quarterly to reflect fluctuation in cases" (Smuck et al., 2011, p. 83).

Sarmento and Silvino (2017) conducted a Portuguese transcultural adaptation and validation study of OPAL. They indicated that because of the significant increase in the growth of clinical research in Brazil the development of indicators that can inform about the quality of work, support evaluation of performance and necessary changes to the existing processes is essential. Due to lack of any instruments to measure workload of clinical research coordinators in Brazil, they aimed to test a transculturally adapted version of OPAL to verify its validity and reliability. Their goal was to be able to better distribute oncology trial protocols among clinical research coordinators that would facilitate them to fulfill the requirements for the trials, as well as identifying the capacity that would require redistribution of protocols and competency training of staff. Their methods included front and back translation of the OPAL tool in Portuguese and conducting a pre-test of the adapted instrument on 15 fictitious protocols in order to define the workload score to set as the gold standard for evaluating the psychometric capacity of the instrument. The final version of the adapted instrument was analyzed to measure the psychometric capacity of the instrument – both reliability and validity of the tool. The numerical data were analyzed using descriptive statistics and concordance of workload scores pertaining to different observed items using inferential statistics, with a 95% confidence interval. Analysis of concordance of observed items included the differences of intra- and inter-observer variability against the average result. The result of their analysis showed a significantly high intra- and inter-observer concordance ($p<0.0001$). The intra-observer analysis resulted in a high level of concordance with ICC scores between 0.987 and 0.934 indicating high reliability. The inter-observer analysis showed high level of concordance compared to the gold standard with an

ICC>0.949 that demonstrated high level of validity of the score. They acknowledged that the absence of a validated tool to calculate workload can result in unrealistic expectations, unmanageable workload and inefficient use of available resources. They concluded that although the adapted OPAL tool met the needs of the users in calculating workload, it is important to be used consistently and that managers allocating protocol assignments evaluate other factors that affect the work of clinical research staff.

In a study conducted by Milani et al. (2017), they tested the application of the Nursing Time Required by Clinical Trial–Assessment Tool (NTRCT-AT). This was a single center study in which the aim was to evaluate this tool as a means to measure workload expressed in time spent to complete trial associated core activities. They identified 30 clinical trial core activities, with 11 related to the trial activation phase and the remainder to study conduct. These 30 core activities were associated with the clinical research nurse's role in coordinating activities of a clinical trial. The NTRCT-AT measured the average times required to complete each of the core activities – expressed as a standard coefficient of time required to perform each activity. NTRCT-AT was used to calculate the total clinical trial nursing time required for coordinating study activities for each study participant enrolled in the trial. They had five phases for study implementation that included the identification of aims, determination of study methods, identification of core activities from the literature and site staff experience, timing of each of the core activities using an objective external observer, data collection and review of the standard co-efficient of time for each activity to calculate average time it took in effort to coordinate the participation of each individual participant. Their study showed that the earlier and more advanced phases of the trials were similar in content – meaning the types of core activities were similar. However, the timing and complexity differed due to the difference in the number of

procedures among various study visit intervals, and the frequency of study visits between different trials. The NTRCT-AT was geared for measuring time commitments of personnel coordinating phase I and III trials. Their data showed that phase I trials required more time for study staff to complete trial activities during the activation phase. However, a weakness of their study was the lack of generalizability of their data since the study was only conducted in one research facility. Nevertheless, their study illustrated that the increasing demands of clinical research necessitated an adequate number of skilled and competent workforce to be involved in each trial.

Good, Lubejko et al. (2013) acknowledged that "[c]oordination of an efficient, successful clinical research program can be challenging" and is compounded by the "lack of resources for quantifying clinical-trial associated workload to help guide staffing and budgetary planning" (p. 211). Therefore, the use of a tool that generates objective metrics to measure workload could potentially increase clinical research program productivity, efficiency and quality of work, while providing managers a guide for calculating staffing and budgetary requirements, as well as evenly distributing work among staff resulting in job satisfaction (Good, Lubejko et al., 2013). They described the evaluation of an acuity-based workload assessment tool developed and implemented to facilitate assessment and balance of workload among research staff. The tool known as the Witchita Community Clinical Oncology Program Acuity Tool (WPAT) was developed by the Witchita Community Clinical Oncology Program (WCCOP). The tool classified study participants into two categories, on a study or off a study. The total number of on- and off-study patients constituted the overall workload. Additionally, protocols were classified as either a treatment- or cancer-control trial and were assigned a score based on six-workload determinants that included treatment complexity, study procedure requirements,

treatment toxicity potential, complexity and number of data forms, degree of coordination required and number of trial random assignments or steps. Development of the tool was based on a patient classification system that was used to determine workload and staffing assignments. The calculation of the acuity scores took place monthly and required multiplying the protocol classification score for each study by the number of active patients from the study that was assigned to staff. Acuity score calculations were conducted over a period of 11 years. Review of the acuity scores over the 11-year period confirmed that the complexity of clinical trials had increased – acuity for treatment trials by 65% from 1999 to 2009 and cancer control by 181% for the same time period (Good, Lubejko et al., 2013). The number of newly enrolled and on-study patients assigned per full-time staff decreased as result of continual monitoring of workload and adjustment of workload based on acuity scores. This demonstrated that "[m]onthly monitoring of individual as well as group average acuity scores provided management with the ability to balance workload among staff" (p. 213).

Good, Hurley et al. (2016) described a project sponsored by the American Society of Clinical Oncology to evaluate the feasibility and utility of a workload measurement tool that applied "objective metrics toward documentation of work, and to provide clearer insight to better meet clinical research program challenges and aid in balancing staff workload" (p. e536). Fifty-one community-based research programs representing 30 states participated in the project, which took place over six consecutive months. According to Good, Hurley et al., a 2010 survey conducted by the ASCO Community Research Forum not only reinforced the need for a method to evaluate clinical trial related workload but was one of the top three of 12 proposed projects. In response, ASCO formed a work group to develop such a tool. After conducting a review of the literature and available tools the work group concluded that three characteristics of a workload

measurement tool should be simplicity, reproducibility and long-term usability. “The project focused only on clinical trial workload associated with patient-centered encounters or clinically focused efforts, defined as any in person protocol-required evaluation and management visit that was designated as required on the protocol study calendar/study plan,” but did not include regulatory-based workload or other non-clinical elements related to clinical trial work (Good, Hurley et al., 2016, p. e537). Two interrelated tools were used for the project: ASCO Protocol Acuity Scoring Worksheet, ASCO Clinical Trial Assessment Tool. The ASCO tool was based on the WPAT (previously described) that had been in use for more than ten years. The acuity scoring worksheet, which consisted of a 4-point rating scale, would be used to calculate the trial complexity level from the least complex to the most complex. The assessment tool was a web-based platform used to collect individual protocol and individual staff acuity scores but was programmed to use full-time equivalent (FTE) status instead of workdays per week used in the WPAT model. The protocol acuity score was multiplied times the number of patient encounters for each specific trial then was divided by the staff FTE value to yield the individual staff acuity score. Descriptive statistics accounted for self-reporting of data by all participating programs, and frequencies and percentages were used to summarize their findings. Additionally, to account for the heterogeneity of the participating program’s characteristics, the programs were grouped into categories to separate them by type and size of the program. Only FTE data were used to summarize staff acuity scores and patient encounters for each group and “data were reported as medians and ranges by patient status, trial sponsor, type of trial and staff title” (Good, Hurley et al., 2016, p. e538). Ninety percent of the 51 research programs that participated provided all six months of workload data, two sites provided five months of data and three sites provided two, three, or four months of data respectively. The research programs reported a median accrual of

150 patients into clinical trials overall and a median of 37 open and actively enrolling trials. They "contributed clinical trial-associated workload data for 323 staff members in total…which represented 963 unique protocols and 165 unique sponsors" (Good, Hurley et al., 2016, p. e539). Six percent of the 323 staff members identified as teams of staff consisting of multiple members. Congruency in the assignment of the same protocol acuity rating was found for 461 protocols but a 1-point difference in 120 protocols and 2-point difference in 23 protocols were found. Variability in assigned acuity rating scores were found in 36% of industry trials, 17% of federally sponsored trials, 3% of academic-sponsored trials and 4% of trials with other sponsors ($P<.001$). The median acuity rating assigned to treatment trials was 3, cancer control trials was 2, correlative science trials was 1.5 and observational/registry trials was 1. Across all groups the highest median staff acuity scores were for those who had patients who were on study and receiving treatment relative to those with patients only on study follow-up. Higher median staff acuity scores were seen in treatment trials, compared to cancer control, observational/registry and prevention trials. Industry trials yielded higher median staff acuity scores than trials that were federally sponsored, academic-sponsored or had other type of sponsors.

According to Coffey et al. (2011) "[t]he increasing use and complexity of multi-modality treatment regimes, the rising costs of clinical trials, the emphasis on the efficient use of available resources and adherence to Good Clinical Practice (GCP), and increasing regulatory requirements and demands for quality assurance/control, have resulted in an increased focus on workload issues" (p. 36). They indicated that the use of unproven methods or simple estimation for measuring workload may result in unrealistic expectations, excessive workload or inefficient use of resources. They described a pilot study of the Workload Measurement Instrument (WMI) that was developed because of the recognition of the importance for understanding the tasks,

time requirements and resources involved in effectively and efficiently conducting clinical research. The WMI "was seen as a means of providing a tool by which individuals could estimate more accurately the time and resources required to participate in clinical trials" (Coffey et al., 2011, p. 36). The development of this instrument involved seven stages leading to the finalization of the trial related tasks to be assessed. The instrument applied concepts from two prior instruments that individually assessed workload and complexity of a trial. The WMI instrument developers divided the instrument into four modules representing the planning, implementation, data management and study close-out phases of a trial. Next, they applied the complexity rating dimension to the instrument. They further subdivided the implementation module into recruitment/treatment and treatment/follow-up to facilitate validating and linking the workload measurement with the complexity dimension. The prospective study only focused on research related activities and not those considered as clinical standard of care. Workload was only assessed for clinical research coordinators to keep participant characteristics homogenous and to facilitate collection of more precise workload data. Data collection was completed over a six-month period. The findings confirmed that all tasks and subtasks within modules 1, 2, 3 were valid, and what tasks or subtasks were completed more and less frequently for modules 1, 2a, 2b and 3. Analysis of the data from modules 4 were not described because of the small number of studies that were in this phase. The aim of this prospective study was focused on validating what trial related tasks both increased the trial's complexity and workload for each of the four phases of the trial conduct included in the modules. The purpose of the WMI was to determine what tasks or subtasks occurred in each of the four phases to allow examination of workload and complexity to inform on multiple different levels.

Several studies or pilot projects involving the testing of a workload measurement tool were described. Of these two were level II high quality quasi-experimental and four were level III good quality non-experimental studies. Despite that some of the pilot testing took place across multiple research facilities, none of their findings were considered generalizable, and emphasis was made for the application of these tools to reflect the heterogeneous characteristics of the research facility to ensure meaningful measurements of workload. Across all the studies, the tools used objective quantitative measurements to calculate trial complexity and workload. The literature suggested that complexity and workload measurement tools are useful in determining workload capacity for different phase trials, estimating staffing and budget requirements and for allocating more balanced workload for research staff. Measurement indicators used to calculate workload are essential in informing about the quality of work, efficient use of staff resources, necessary changes to existing processes, competency training related to trial task requirements and for supporting the evaluation of clinical trial performance. The literature confirms phase I to be the most complex level of trials.

## Theoretical Framework

Two conceptual models were utilized as the framework for the development and implementation of this pilot project. The Institute for Healthcare Improvement (IHI) Model was used to guide the development of the project aims. This model incorporates the use of Plan-Do-Study-Act (PDSA) cycle for testing and studying the change for improvement (Appendix C).

### IHI Model for Improvement

The three fundamental questions from the IHI model guided the development of the project aims, measures to determine if change leads to improvements and changes to implement that result in process improvement (Institute for Healthcare Improvements, n.d.). Utilizing the

PDSA cycle, the IHI model outlined four main components for the iterative process for testing and studying the change – the use of a quantitative workload measurement tool to assess trial complexity and workload.

- Step 1 (Plan) involves planning of testing and observation that includes a data collection plan
- Step 2 (Do) involves small-scale testing of an intervention or change in process
- Step 3 (Study) involves time to analyze the data and study the results
- Step 4 (Act) involves making refinements to the intervention or process change based on information learned from the test (Institute for Healthcare Improvements, n.d.).

**Plan-Do-Study-Act**

According to Coury et al. (2017), the PDSA cycle is a commonly used model for implementing small tests of change to optimize process improvement in health care settings that might have untapped potential for pragmatic research. In contrast to clinical trials which "emphasize internal rather external validity, using highly controlled environments and selected populations…pragmatic studies are generally embedded in care delivery environments," therefore making this an ideal model to use as the strategy for implementation of this pilot project (p. 2). Its advantage is it allows evaluation and refinements of an intervention to be made soon after it is implemented.

Taylor et al. (2014) cited that there is mixed evidence on effective quality improvement interventions and many have concluded that such interventions may only be effective in specific settings. Additionally, they indicated that research findings emphasized the effect that the local context has on the success of an intervention, citing that stand-alone interventions are unlikely to deliver consistent improvements and instead require multi-faceted approaches that are developed

iteratively to adapt to the predictable and unpredictable changes that occur over time and within a complex social system. They further cited that the PDSA cycle is one such method for an iterative development of change, either used by itself or as part of a wider quality improvement approach, such as the IHI Model for Improvement (Taylor et al., 2014).

## Methodology

Using the PDSA cycle, project implementation was divided into two stages. The first stage was project planning and the second was the implementation stage. During the planning stage, the project stakeholders were identified, the aims and objectives were developed, stakeholder roles, responsibilities and project timelines were established, the data collection plan was created, workload measurements for two sample protocols were performed and adjustments were made to the data collection plan and elements of the tool based on the initial measurements performed.

The following were completed during the implementation stage: measurement of trial complexity and workload was completed for ten protocols that were active between 2011 to 2019 inclusive, a job stress survey of the clinical trials unit personnel, and a review of protocol deviations reported for the ten protocols from the project sample.

### Project Setting & Population

The pilot project was conducted on a moderate sized clinical trials unit located within an academic and tertiary healthcare institution. The unit consisted of three research managers, five clinical research nurses (CRN) and ten clinical research coordinators (CRC). The research staff is responsible for coordinating trials that were mostly phase I and investigating the use of cell-based immunotherapies to treat patients with advanced stage cancer. Trial conduct involved physician investigators, multidisciplinary clinical teams, study sponsor teams (e.g., project and

data management), and three institutional oversight committees. Technical assistance was solicited from the institution's Quality Improvement consultant for advice on certain elements related to the project, such as Lean principles and how to perform time measurements of work activities. A staff statistician was also consulted for advice on data collection and interpretation.

**Project Feasibility**

The Institutional Review Board of record (i.e., University IRB) granted an exemption for this quality improvement (QI) project. The project site's IRB deferred to the IRB of record for the exemption for this QI project. The department of nursing education for the project site concurred with the project site's IRB and permitted the project to proceed as planned. Additionally, the project aims and its implementation received the support from the project lead's executive deputy director who functioned as the residency preceptor and the personnel from the clinical trials unit who were stakeholders for the QI project implementation.

Minimal expense was incurred for this pilot project. All of the data collection worksheets and instruction guides were created with Microsoft (MS) Word, and data were entered into MS Excel for the eventual data review and interpretation. This software application was a work issued tool that was available to the project lead, research managers and research staff prior to the start of the project and therefore did not generate any additional expenses.

There were some anticipated barriers for the implementation of this pilot project. The integration of project activities into the existing workflow of the research managers and staff (i.e., CRNs, CRCs) who had many competing work priorities was a concern. As stakeholders they participated in the development of the data collection tools and participated in data collection and other project related activities. These barriers were mitigated with scheduled project meetings and use of data collection tools that could be completed electronically and

remotely during the approximately two-month period when the data collection occurred. The completion of data collection was not extremely time consuming for the stakeholders nor were these extremely time sensitive. The data collection activities for this project were completed in approximately two months with all of the actual data collection being completed remotely. The research managers and staff had access to the unit's trial master files and share drive via secure remote access that facilitated their ability to complete their administrative work duties and project related data collection off site.

The other anticipated barriers were confidentiality and intellectual property concerns. Confidentiality considerations pertained to the staff who participated in the pilot project, as well as with the use of study participant information. To maintain anonymity of staff and study participants only deidentified data were used for data collection. The project site's legal representative confirmed that there was no intellectual property issue since this was not a research project.

There was also an ethical concern related to the potential risk for bias because of the dual role I had as the director of the unit and graduate student conducting the QI project. It was crucial that my role as the director did not intersect with my role as the project lead. This was of primary concern with the workload survey category pertaining to job satisfaction. To reduce bias, stakeholders were involved with the development of the data collection tools pertaining to workload measurement to ensure that it represented the specific needs of their unit. Objective tools such as OPAL and an established job stress survey available in the public domain were used to collect data for workload measurements and job stress factors. With respect to the survey, steps were taken to preserve the anonymity of the respondents by having them return their completed questionnaire via email. The questionnaire which did not include any identifiers

were then immediately saved into an electronic folder and the email was then deleted. The review of the responses took place after all of the completed questionnaires were filed. The demographic questions included in the survey pertained to educational background and research experience and did not include questions that would likely disclose the identity of the respondents.

The COVID-19 pandemic occurred at the onset and persisted throughout the duration of the data collection for this project. Due to many significant changes to the institution's clinical workflows resulting from the pandemic and social distancing requirements, the time study component for this pilot project was permanently eliminated. Other components for the project were unaffected by the COVID-19 crisis.

**Trial Complexity & Workload Measurement**

A comparison was made between the various instruments studied to measure trial complexity and workload in the clinical trial setting from the description in the literature. OPAL was selected as the tool for this pilot project based on the guiding principles that it was easy to use, adaptable to the local context of the clinical trials unit where the project took place, and measured trial complexity. Three measurements were generated using OPAL: trial complexity, case workload, and total workload. Workload measurements were performed by two of the unit's research managers for ten clinical trial protocols using OPAL. They collaborated with the project lead in recreating the list of central processes (CP) and special procedures (SP) to reflect the trials in our unit (Appendix B). They performed initial measurements for two trial protocols, then met with the project lead to discuss modifications necessary for clarifications for the calculation of case workload. Refinements to both the workload measurement worksheet and instruction guide were made to facilitate this need.

A trial complexity score was assigned to each of the ten protocols using the OPAL pyramid scale (Appendix D). The pyramid scale complexity level is based on the phase of the trial and a set of criteria that include the number of occurrences of SPs and CPs. We modified the list of CPs and SPs to make it more representative of the trial activities for our unit.

The original calculation for case workload involved multiplying the OPAL complexity score by the number of study participants in either the active or follow-up phase of the trial. Calculation of the total workload involved the addition of the OPAL complexity score plus the case workload score. The case workload score represented patient management and the total workload score was associated with study management. For the convenience of the pilot testing of OPAL we multiplied the complexity by only the number of patients for the trial sample size.

We were interested in learning if we could better assess the trial complexity by adding points for every occurrence of a SP that we identified as unique to certain protocols (Appendix B). We modified the calculation of case workload by adding the complexity score plus the additional points for SP and multiplied this by the number of patients for the trial sample size yielding new case workload scores for the ten protocols in our sample.

**Protocol Deviations**

In parallel, additional stakeholders were also solicited to participate in data collection of protocol deviations that were reported for each of the ten trial protocols from our sample. An index of all protocol deviations was available from each protocol's trial master file. The regulatory manager from the unit created copies of the tables that indexed all of the protocol deviations that were reported to date. Review of protocol deviations was an important concept for examining errors made during protocol execution that would inform about factors affecting

accuracy, competency of staff and factors that contribute to work stress. We also wanted to learn if any of the deviations impacted patient safety or study outcome.

**Job Stress Survey**

A workload survey was assembled using four categories from a job stress questionnaire created by the National Institute for Occupational Safety and Health (NIOSH) that was available in the public domain. The NIOSH Generic Job Stress Questionnaire is a validated instrument that included multiple dimensions of job stress that could be used in total or individually to construct a separate survey questionnaire. Four categories from this instrument were used to develop the workload survey for this project.

- Job Requirements
- Workload and Responsibility
- Mental Demands
- Job Satisfaction

Questions from the NIOSH instrument representing these dimensions were included in the survey. The intent of the survey was to examine if there were any correlations between the responses to the survey and factors impacting workload. In particular, this was related to the impact on staff efficiency and work stress. Instructions accompanying the survey questionnaire was emailed to 18 staff members of the clinical trials unit. The questionnaire was designed to be completed electronically and returned to the project lead via email. The questionnaire did not require any identifiers that would discourage honest responses to each of the questions, most especially those pertaining to job satisfaction. A minimal amount of demographic information pertaining to educational background and research experience was queried on the survey but none that would likely compromise the anonymity of the respondents. The participants were

requested to return their completed survey form via email and each form was blindly copied from the email into an electronic folder and were assigned a file number based on the sequential order in which these were received. This was to keep accounting of responses for follow-up purposes. The email from the participant was deleted after their completed questionnaire was filed. All correspondences related to the survey were conducted using the project lead's school email account.

## Outcomes Measurement & Findings

The outcomes and interpretation of the findings were guided by Lean principles. According to Sweeney (2017), Lean incorporates the "philosophy of kaizen, or the creation of a culture of continuous improvement" (p. 9). In measuring the outcomes of our aims and interpreting the findings we utilize the concepts of Muda, Mura and Muri from Lean to focus our attention on the opportunities for process improvement changes. Each of these concepts of Lean addresses fundamental areas for process improvements. "Muda represents waste in its most physical form," and "[t]he objective with waste reduction and elimination is to clearly separate the value-added activities that are identified as wasteful or non-value-added" (Sweeney, 2017, p. 7). "Mura is waste in the sense of unevenness," because "[u]nevenness in workflow can result in unnecessary downtime or periods of unnecessary stress on equipment, systems, and workforce" (Sweeney, 2017, p. 8). "Muri is a failure to understand capabilities or to succumb to the effects of overburden" (Sweeney, 2017, p. 8).

### Objective 1/Outcome 1

The first thing we were trying to accomplish was to show that we could measure trial complexity. We used OPAL to measure complexity for ten trial protocols that were conducted between 2011 and 2019 inclusive. What we learned was that we could determine trial complexity

using OPAL. Of our ten trial protocols, six were phase I, one was phase I/II and three were phase II. All phase I and I/II trials scored a complexity level of eight (8) that was consistent with the literature noting phase I trials as the most complex. It is also notable that most of the trials conducted in this research unit are phase I trials.

**Objective 1/Outcome 2**

We wanted to see if the number and frequency of trial procedures affected the complexity and workload scores. We completed two calculations for case workload using the original and a modified formula to generate original and new workload scores. What we found was that the new workload scores were greater than the original case workload scores with a mean of 416.5 and 173.2 respectively (Table 1). The modified calculation for case workload then resulted in higher total workload scores.

**Table 1**

*Case Workload Scores*

| | OPAL | Original Case Workload | Original Total Workload | New Case Workload | New Total Workload |
|---|---|---|---|---|---|
| Mean | 7.70 | **173.20** | 180.90 | **416.50** | 424.20 |
| Median | 8.00 | 164.00 | 172.00 | 419.00 | 427.00 |
| SD | .483 | 54.121 | 54.098 | 178.727 | 178.854 |
| Minimum | 7 | 98 | 105 | 201 | 208 |
| Maximum | 8 | 264 | 272 | 792 | 800 |

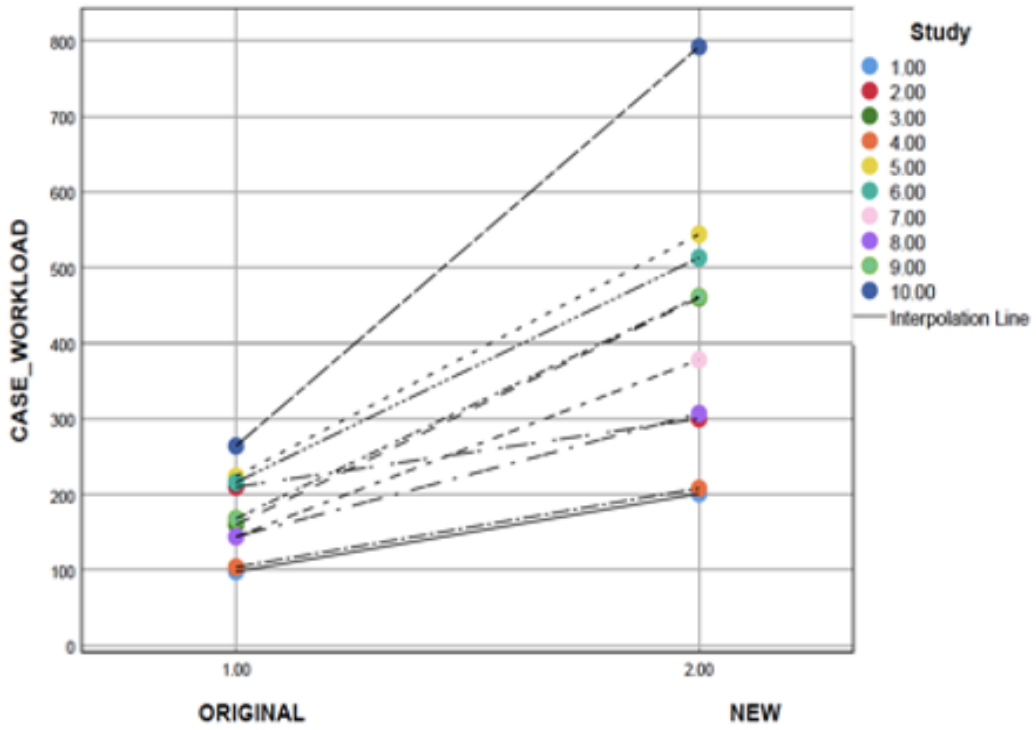
We measured the difference between the original and new case workload scores using a paired t-test. To illustrate the difference, we plotted that scores on a graph (Figure 1). The paired t-test strongly suggests a highly significant difference in the mean values of the two scores with a $P<.001$. This suggested that the new case workload scores captured the true complexity of these trials.

**Objective 1/Outcome 3**

We were also interested in learning if we could create a threshold representing an acceptable distribution of work. We plotted the mean case workload scores and values of 1 SD and 2 SD on a graph (Figure 2). Using a random selection of two case workload scores from our sample of ten trials we calculated the workload scores for four simulated CRCs and plotted this on the graph to see whether the simulated workload scores were extreme (above 1 SD) or acceptable (within 1 SD). If scores lie within 1 SD of the mean this suggested an acceptable distribution of work, while those above would be considered extreme. To note, since our project sample only contained ten trial protocols the scores were not precise so measurements should be

**Figure 1**

*Original & New Case Workload Scores*




*Note.* Y-axis shows the range of case workload scores. X-axis shows the original and new case workload scores for all ten trial protocols from our project sample.

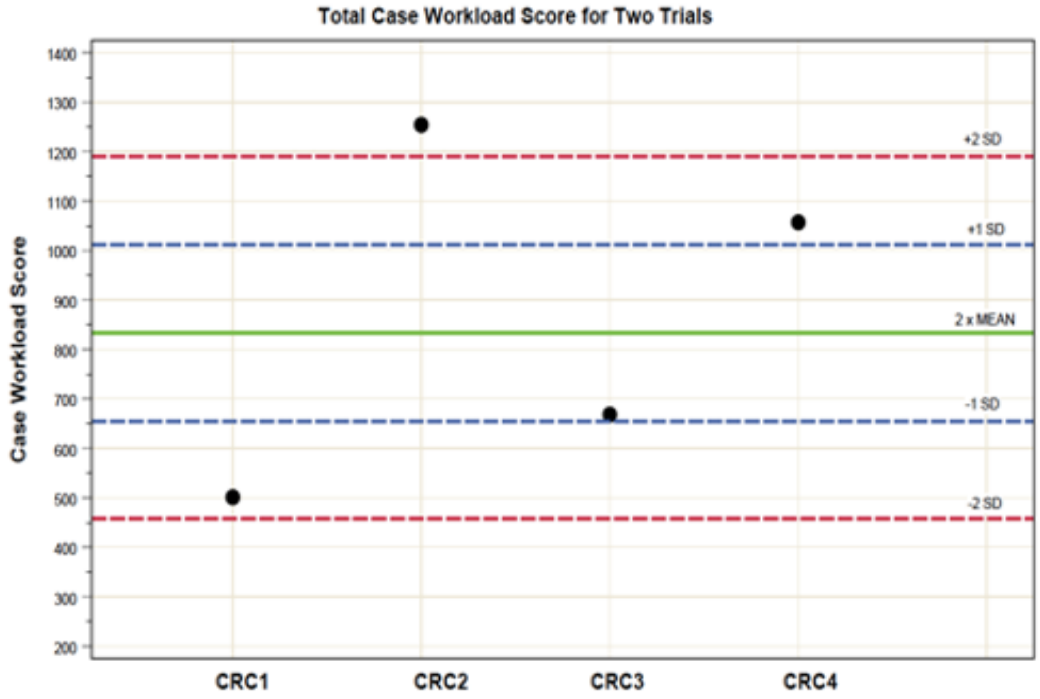
replicated using a larger sample. Nevertheless, we were able to obtain useful information to inform on a balanced distribution of work.

**Objective 2 Outcome**

We wanted to identify three factors that impact staff workload. When we examined the original case workload against the new workload scores (Figure 1), we saw a significant increase in the scores, on the average a 2.3-fold difference suggesting the volume and frequency of trial procedures likely contribute to increase in workload. From our paired t-test complexity of trial is another factor.

**Figure 2**

*Workload Threshold*




*Note.* Y-axis represents the range of case workload scores. X-axis represents the clinical research coordinators (CRC) simulated case workload scores.

**Objective 3 Outcome**

We wanted to identify four factors that affect efficiency and contribute to job stress. We conducted a workload survey that collected a limited amount of demographic data pertaining to staff educational background and work experience. The largest proportion of staff had either a healthcare or science related background. Of these two groups, most of the staff with healthcare background had prior jobs in clinical research and 6 to 10 or greater years of experience in clinical research.

Below were the more frequent responses from the job requirement category of the survey (Appendix E). The more extreme of these responses are in bold. We interpreted these responses to mean that although their job was challenging, their skills were often matched with the tasks they needed to complete, suggesting that skill level should match the complexity of the work tasks to be efficient.

- ~ 47% = required to work fast sometimes
- **~43% = required to work hard fairly often**
- ~37% = had little time to get things done sometimes
- ~41% = had a great deal to be done sometimes
- **~41% = had a great deal to be done fairly often**
- ~58% = had marked increase in workload sometimes
- **~43% = had marked increase in amount of concentration required on the job fairly often**
- **~43% = were given a chance to do the things they do best fairly often**
- **50% = used skills from previous experience and training fairly often**

The responses from the workload and responsibility category (Appendix F) indicated that although ~62% of the staff in this unit had a lot of work, 50% of them had a lot of time to complete their work suggesting that there is an even distribution of work. Fifty-nine percent indicated having a lot of projects, assignments or tasks suggesting that this may be an area requiring further examination of skill-mix.

- 50% = experience a little slowdown in workload
- ~68% = have some time to think and contemplate
- ~62% = have a lot of workload
- 50% = have a lot of time to do all their work
- ~59% = have a lot of projects, assignments or tasks
- ~56% = have some lulls between heavy workload periods

From the mental demands category (Appendix G), approximately 68% indicated that their job required a great deal of concentration and about 80% were required to remember many different things. We interpreted this to mean that although their workload is likely balanced their work is very challenging and complex.

With respect to job satisfaction (Appendix H), 50% indicated they were very satisfied and the other 50% indicated they were somewhat satisfied. Approximately 87% indicated that they would take the same job over again and 68% would choose a similar job if they had other choices. This suggested that their current work and workload were not affecting overall job satisfaction.

When we reviewed the protocol deviations that were reported for four of the ten protocols from our project sample, we noted that the higher numbers of deviations pertained to missed tests or procedures and study visits or procedures completed outside of the protocol specified window.

This suggested that the volume and frequency of protocol required tests and procedures affected the accuracy and efficiency of the staff's work performance. This is consistent with the literature that indicated that humans who are overburdened with work may hurry their performance, commit more errors and have poor awareness of their surroundings. It is notable that none of these protocol deviations were determined to affect patient safety or study outcome.

**Objective 4 Outcome**

What we learned from our findings is that the volume, frequency and complexity of trial work activities does affect the clinical trial performance in our unit. This was evident from our workload assessments, as well as the number of protocol deviations that we noted from our review. We also learned that the volume of work, complexity of trial related tasks, aggressive timelines, prior work experience, skill level and strategic use of staff skill mix should be factors to assess because they do impact staff efficiency and can contribute to work stress. For example, matching skill level to the complexity of the task required for staff to complete would result in a more efficient use of staff resources. We also learned that there is value in using quantitative measurements to assess workload. More importantly we learned that use of established conceptual models such as the IHI Model for Improvement, PDSA and Lean are more reliable and sustainable methods for testing and studying change directed at process improvements.

As a result of our pilot project, we are piloting the use of a staff effort calculator that we created using MS Excel. The calculator is used to calculate staff effort dedicated to each of their assigned protocols based on hours worked in a 40-hour week. This calculator can also be used to perform a cost benefit analysis for staffing requirements based on the skill level required for the role and responsibilities for each trial protocol.

## Limitations

The overall limitation for this pilot project was the limited sample size of the protocols used for measuring trial complexity and workload. We also did not calculate case workload using active and follow-up patients actually enrolled in the trials. Another limitation was that we only measured trial complexity but not the complexity of the individual tasks themselves. This would be informative in determining whether volume of work, the complexity of the task or both, impact clinical trial performance and efficiency. We also had time restrictions for completing repeat testing of the tool and performing additional measurements. Finally, due to concerns of confidentiality the responses to the survey pertaining to job satisfaction may not have been as candid.

## Conclusion

The most important lesson we learned from this pilot project was that there are established and more exquisite methods for testing and studying changes for process improvement. I recommend that concepts from the IHI Model for Improvement, PDSA and Lean be used in this unit for implementation and evaluation of process improvement changes as these are more sustainable comprehensive methods in comparison to single-bullet approaches. Consistent evaluation of our operational model needs to be conducted to determine when process improvement changes are warranted. Use of quantitative methods to calculate staff effort and work capacity will result in more accurate estimation of staffing and budgetary requirements.

## References

Brennan, C. W., Krumlauf, M., Feigenbaum, K., Gartrekk, K., & Cusack, G. (2019). Patient acuity related to clinical research: Concept clarification and literature review. *Western Journal of Nursing Research*, *41*(9), 1306–1331. https://doi.org/10.1177/0193945918804545

Casner, S. M., & Gore, B. F. (July 2010). *Measuring and evaluating workload: A primer* (NASA/TM—2010-216395). NASA Center for Aerospace Information. https://doi.org/https://matb-files.larc.nasa.gov/Workload_Primer_TM_Final.pdf

Coffey, M., Berridge, J., Lyddiard, J., & Briggs, J. (2011). Workload measurement instrument. *Applied Clinical Trials*, *20*(1), 36–40. https://doi.org/http://www.appliedclinicaltrialsonline.com/workload-measurement-instrument

Coury, J., Schneider, J. L., Rivelli, J. S., Petrik, A. F., Seibel, E., D'Agostini, B., Taplin, S. H., Green, B. B., & Coronado, G. D. (2017). Applying the Plan-Do-Study-Act (PDSA) approach to a large pragmatic study involving safety net clinics. *BMC Health Services Research*, *17*(411), 1–10. https://doi.org/10.1186/s12913-017-2364-3

Dang, D., & Dearholt, S. L. (2018). *Johns Hopkins nursing evidence-based practice: Model and guidelines* (3rd ed.). Sigma Theta Tau International .

Getz, K. A., Wenger, J., Campo, R. A., Seguine, E. S., & Kaitin, K. I. (2008). Assessing the impact of protocol design changes on clinical trial performance. *American Journal of Therapeutics*, *15*, 450–457. https://doi.org/doi: 10.1097/MJT.0b013e31816b9027

Good, M. J., Hurley, P., Woo, K. M., Szczepanek, C., Stewart, T., Robert, N., Lyss, A., Gonen, M., & Lilenbaum, R. (2016). Assessing clinical trial-associated workload in community-

based research programs using the ASCO clinical trial workload assessment tool. *Journal of Oncology Practice*, *12*(5), e536–e547. https://doi.org/http://jop.ascopubs.org/lookup/doi/10.1200/JOP.2015.008920

Good, M. J., Lubejko, B., Humphries, K., & Medders, A. (2013). Measuring clinical trial-associated workload in a community clinical oncology program. *Journal of Oncology Practice*, *9*(4), 211–215. https://doi.org/10.1200/JOP.2012.000797

Institute for Healthcare Improvement. (n.d.). *Science of improvement: How to improve*. Retrieved April 27, 2020, from http://www.ihi.org/resources/Pages/HowtoImprove/ScienceofImprovementHowtoImprove.aspx

Institute for Healthcare Improvement. (n.d.). *Science of improvement: Testing changes*. Retrieved April 27, 2020, from http://www.ihi.org/resources/Pages/HowtoImprove/ScienceofImprovementTestingChanges.aspx

Malik, L., & Lu, D. (2019). Increasing complexity in oncology phase I clinical trials. *Investigational New Drugs*, *37*, 519–523. https://doi.org/10.1007/s10637-018-0699-1

Milani, A., Mazzocco, K., Stucchi, S., Magon, G., Pravettoni, G., Passoni, C., Ciccarelli, C., Tonali, A., Profeta, T., & Saiani, L. (2017). How many research nurses for how many clinical trials in an oncology setting? Definition of the Nursing Time Required by Clinical Trial-Assessment Tool (NTRCT-AT). *International Journal of Nursing Practice*, *23*, 1–8. https://doi.org/10.1111/ijn.12497National Institute for Occupational Safety and Health. (n.d.). *NIOSH Generic Job Stress Questionnaire*. https://www.cdc.gov/niosh/topics/workorg/detail088.html

National Institute for Occupational Safety and Health. (n.d.). *NIOSH generic job stress questionnaire*. https://doi.org/https://www.cdc.gov/niosh/topics/workorg/detail088.html

Sarmento, R. M., & Silvino, Z. R. (2017). Measuring workload of clinical trials: Transcultural adaptation and validation to Portuguese language of Ontario Protocol Assessment Level (OPAL). *Journal of Clinical Research & Bioethics*, *8*(4), 1–6. https://doi.org/10.4172/2155-9627.1000308

Smuck, B., Bettello, P., Berghout, K., Hanna, T., Kowaleski, B., Phippard, L., Au, D., & Friel, K. (2011). Ontario protocol assessment level: Clinical trial complexity rating tool for workload planning in oncology clinical trials. *Journal of Oncology Practice*, *7*(2), 80–84. https://doi.org/doi:10.1200/JOP.2010.000051

Sweeney, B. (2017). *The simplified beginner's guide to Lean* (2nd ed.). Clyde Bank Media LLC.

Taylor, M. J., McNicholas, C., Nicolay, C., Darzi, A., Bell, D., & Reed, J. E. (2014). Systematic review of the application of the plan-do-study-act method to improve quality in healthcare. *British Medical Journal Quality & Safety*, *23*, 290–298. https://doi.org/10.1136/bmjqs-2013-001862

# Appendices

## Appendix A

*Key Literature Synthesis and Appraisal*

| Author | Title | Synthesis | Findings | Appraisal |
|---|---|---|---|---|
| Smuck et al. (2011) | Ontario protocol assessment level: Clinical trial complexity rating tool for workload planning in oncology clinical trials. | • Aim - to demonstrate that OPAL was reliable, and scoring was consistent across clinical trial protocols and trials sites<br>• OPAL workload measurement tool that rated protocol complexity based on a pyramid scale and calculated workload scores<br>• Tool developed by experience group of clinical trial managers<br>• Guiding principles for tool development were – easy to apply, measured work of clinical trial professionals, useful for calculating workload capacity, included all trial phases, and not intended for academic review<br>• Complexity levels based on trial phase and protocol activities – central processes and special procedures<br>• Calculation of workload score based on OPAL complexity level<br>• Tool tested in 17 participating cancer clinical trials sites with heterogenous characteristics<br>• 176 OPAL assessments were completed for 27 protocols | • From the 176 assessments completed, analysis of the score variance of 1.5 was considered acceptable on the basis of research practice models and participation in correlative studies<br>• Two trials were rated with a variance of 3.5<br>• Misinterpretation of the definition of special procedures produced higher OPAL scores that indicated that OPAL needs to be applied consistently at the site and be based on local practices to produce measurable site workload information | Level III – Good Quality Non-experimental study |

| Author | Title | Synthesis | Findings | Appraisal |
|---|---|---|---|---|
| Sarmento & Silvino (2017) | Measuring workload of clinical trials: Transcultural adaptation and validation to portuguese language of Ontario protocol assessment level (opal). | • Aim – to perform Portuguese transcultural adaptation and test its validity and reliability<br>• OPAL tool translated to Portuguese<br>• Pretest of translated tool performed on 15 fictitious protocols to define workload score to set as the gold standard for evaluating psychometric capacity of the tool<br>• Final version tested for reliability and validity<br>• Numerical data analyzed using descriptive statistics<br>• Concordance of workload scores pertaining to different observed items analyzed using inferential statistics<br>• Confidence interval 95% | • Analysis of concordance of observed items included differences of intra- and inter-observer variability against the average result<br>• High intra- and inter-observer concordance (p,0.0001)<br>• High level of intra-observer concordance with ICC scores between 0.987 and 0.934<br>• High level of inter-observer concordance compared to the gold standard score with an ICC.0.949 demonstrating high level of validity of the score.<br>• Concluded after analysis that adapted tool met the needs of the users in calculating workload but needs to be used consistently and other factors that affect work must also be evaluated | Level II – High Quality Quasi-experimental study |
| Milani et al. (2017) | How many research nurses for how many clinical trials in an oncology setting? Definition of the Nursing Time Required by Clinical Trial-Assessment Tool (NTRCT- AT). | • Aim – to evaluate NTRCT-AT as a means to measure workload expressed in time spent to complete trial associated core activities<br>• Time measurements used to calculate total clinical trials nursing time required per patient enrolled in a trial<br>• Participants comprised of 7 Italian clinical research nurses<br>• Research nurse team identified aim and study methods<br>• A list of 30 core activities were selected for time measurements – 11 pertaining to study activation and the remaining to study conduct<br>• Time measurements of core activities performed by independent observer<br>• Pilot testing conducted in one research facility<br>• All activities were timed for all research nurses to reduce individual variability | • Compared the total nursing time required by the trials annually (1,254, 578 minutes/year) with the total number of working days required in theory and the actual worked days showed a greater theoretical workload in hours per year (average 11.13 hours for one year assessed)<br>• The research nurse participants felt this to be an accurate reflection of their current situation of excessive workload and difficulty in carrying out all necessary activities | Level III – Good Quality Nonexperimental study |

| Author | Title | Synthesis | Findings | Appraisal |
|---|---|---|---|---|
| | | • Tool tested retrospectively on 141 clinical trials<br>• Provides a quantitative measurement of workload that trial work activities represent in a workday | | |
| Good et al. (2013) | Measuring clinical trial-associated workload in a community clinical oncology program. | • Aim – to evaluate an acuity-based workload measurement tool (WPAT) developed by WCCOP<br>• Patient Classification - classified trial participants into two categories – on-study or off-study<br>• Total number of on- and off-study participants constituted the overall workload<br>• On-study group subcategorized into active treatment and off-treatment<br>• Protocol Classification – protocol classified as either treatment-focused or cancer control-focused and ranked based on 6 workload-related determinants<br>• Acuity scores assigned to individual clinical trials according to their estimated workload ranging from 1-observational, 2-requiring oral agents with minimal toxicity potential, 3-chemotherapy with increased toxicity potential, complex drug regimens high toxicity potentials<br>• Calculated individual research nurse workload scores monthly for 11 years by multiplying number of patients per trial by the assigned acuity score assigned to trial | • Review of acuity score data across an 11-year timespan confirmed clinical trial complexity had increased<br>• Treat-focused trials acuity scores increased by an annual average of 65% versus cancer control-focused trials by 181% from 1999 – 2009<br>• Treatment trial 11-year average acuity score was 30.6, with annual average ranging from 19.3 to 45.6 between 2000 – 2008<br>• Cancer Control trial annual acuity increased from 8.97 to 69.8 from 2001-2006<br>• Off-treatment acuity scores showed a 178% increase from 2001-2009<br>• The number of new patient enrollment and the number of patients on and off study increased over the 11-year timespan but as a result of monitoring workload and adjusting staffing needs based on acuity scores, the number of patients per FTE research nurse for new enrollments and patients on study decreased. The number of patients categorized as off study only slightly increased. | Level III – Good Quality Nonexperimental study |

| Author | Title | Synthesis | Findings | Appraisal |
|---|---|---|---|---|
| Good et al. (2016) | Assessing clinical trial-associated workload in community-based research programs using the ASCO clinical trial workload assessment tool. | • Aims – to test a combination to two interrelated tools measuring protocol complexity and workload effort for clinical trials across multiple practice settings known as Clinical Trial Workload Assessment Tool<br>• Two interrelated tools: ASCO Protocol Acuity Scoring Worksheet incorporated a 4-point protocol complexity rating scale where score of 1.0 = lower complexity/workload to 4.0 = complex trial with greater workload. ASCO Clinical Trial Workload Assessment Tool a web-based platform facilitated collection of clinical trial-associated workload data.<br>• 51 community-based research programs representing 30 states participated in this project<br>• Clinical trial-associated monthly workload data were collected and entered into the web-based tool for 6 consecutive months<br>• Project tool used refined and edited WPAT 4-point protocol scoring criteria developed previously by WCCOP<br>• Protocol complexity assessed before assignment of protocol acuity score<br>• Project tool accounted for two acuity metrics – protocol acuity and individual staff acuity scores<br>• FTE status was programmed to be used instead of days worked per week<br>• Patient encounters recorded in tool as either on-study but off treatment or off-study in follow-up<br>• Calculation of individual staff acuity score: number of patient encounters x protocol acuity score/staff member FTE value = individual staff acuity score | • Descriptive statistics were computed for self-reported program characteristics and workload data, including staff acuity scores and number of patients encounters<br>• Self-reported characteristics of the programs revealed a variety of types of programs, various degrees of experience and accrual volumes<br>• 47% = federally funded<br>• 14% = community hospital<br>• 14% = non-academic hospital-based private physician practice<br>• 22% = non-hospital based private physician practice<br>• 2% = private research network<br>• 2% = other<br>• Clinical trials experience ranged from 7 to 30 years<br>• Clinical trial accrual median = 150 patients<br>• Open trial median = 37<br>• Actively accruing trials ranging from 9 to 186<br>• Research programs contributed clinical trial-associated workload data for 323 staff members, including RNs (49%), CRAs (28%), research coordinators (16%), administrators/managers (1%)<br>• Data represented 963 unique protocols of which 604 observed patients on study treatment<br>• Data represented 165 unique sponsors<br>• 6% of staff identified as “teams of staff that included multiple members<br>• Participating program were shown to be congruent in assignment of same protocol acuity rating to 461 protocols with patients on study treatment (76%) | Level II – High Quality Quasi-experimental study |

| Author | Title | Synthesis | Findings | Appraisal |
|---|---|---|---|---|
| | | • Project only focused on core and consistent elements rather than all elements of the clinical trial-associated workload | • A difference of one point was found in 120 protocols (20%)<br>• Two-point difference reported in 23 protocols (4%)<br>• Variability in assigned scores found in 36% on industry trials, compared with only 17% of federally sponsored trial, 3% academic-sponsored trials and 4% of trials with other sponsors (P,.001).<br>• Median acuity ratings assigned to treatment trials was 3; cancer control trials assigned median of 2, correlative science trials assigned a median of 1.5 and observational/registry trials a median of 1<br>• 96% of the 51 participating programs provided at least 5 months<br>• Response rate along with feedback form participants demonstrated that the tool was simple and easy to use and supported long-term feasibility and utility for community-based research programs and also minimized bias in the findings<br>• Results support the idea that complexity of trial affects the work associated with trials<br>• Establishing single-benchmark acuity score associated with number of patient encounters for FTE staff member to use as reference for community-based research programs required grouping programs under similar categories to allow comparison to the most applicable or similar types of program<br>• Provides preliminary understanding of the complexity of the measurement of clinical trial-associated workload | |
| Coffey et al. (2011) | Workload measurement instrument. | • Aim – to collect data using Workload Measurement | • Total of 414 modules were completed from 27 hospitals | Level III – Good Quality Non-experimental study |

| Author | Title | Synthesis | Findings | Appraisal |
|---|---|---|---|---|
| | | Instrument (WMI) to validate the four WMI modules linked to a previously developed trial complexity assessment tool<br>• WMI developed in 7 stages: 1-development, 2-validation, 3-revision of draft check list of trial related activities, 4-drafting of the WMI, 5-feasibility pilot study, 6-analysis, 7-revision of trial related tasks incorporated into the 4 modules<br>• Prospective study focused only on research related activity and not those considered routine care<br>• Workload only recorded for research staff working on trial and excluded investigators, pharmacists, day care staff, etc. due to belief that it would be difficult to collect accurate data from such a diverse range of staff – accepted limitation of study<br>• Four WMI modules: Module 1-planning stage, Module 2-implementation stage, Module 3-trial data management stage, Module 4-closure/final stage<br>• Module 2 further subdivided in 2a-recruitment and 2b-treatment/follow-up<br>• Data collection occurred over a 6-month period<br>• Modules 1 and 4 included any trials in planning and closure stage due to the small number of trials likely to be these stages during the 6-month data collection period<br>• Modules 2 and 3 included trials covering as many tumor types, treatment modalities and study types to ensure an inclusive range of study activities<br>• 36 studies were included in the prospective study | covering 12 UK cancer research networks<br>• Data completed for 35 of the 36 trials included in study<br>• Three centers did not complete the modules due to merger of networks, staff shortage or excessive workload issues<br>• Assumption were made that research staff were experienced and trained; not involved in actual delivery of treatment; research activity was additional to standard treatment; hospital support services, standard equipment was available/accessible to staff<br>• Module 1 = most frequently completed subtasks were those related to preparation/submission of documentation to ethics committees; lowest recorded tasks related to organizing and attending meetings other than in-house meeting for information, training, and trial activation<br>• Module 2a = most frequently completed subtasks related to screening for eligibility, informed consent, and sample preparation – consistent with expectations and personal experience of the project participants; least completed subtask was coordinating and verifying radiotherapy dose reductions and missed treatment – anticipated given very low numbers of radiotherapy trials included in the study<br>• Modules 2a, 3 = most frequently completed subtask was completion of case report form, photocopying/sending trial documents, preparation and submission of trial amendments and updating study documents following amendments | |

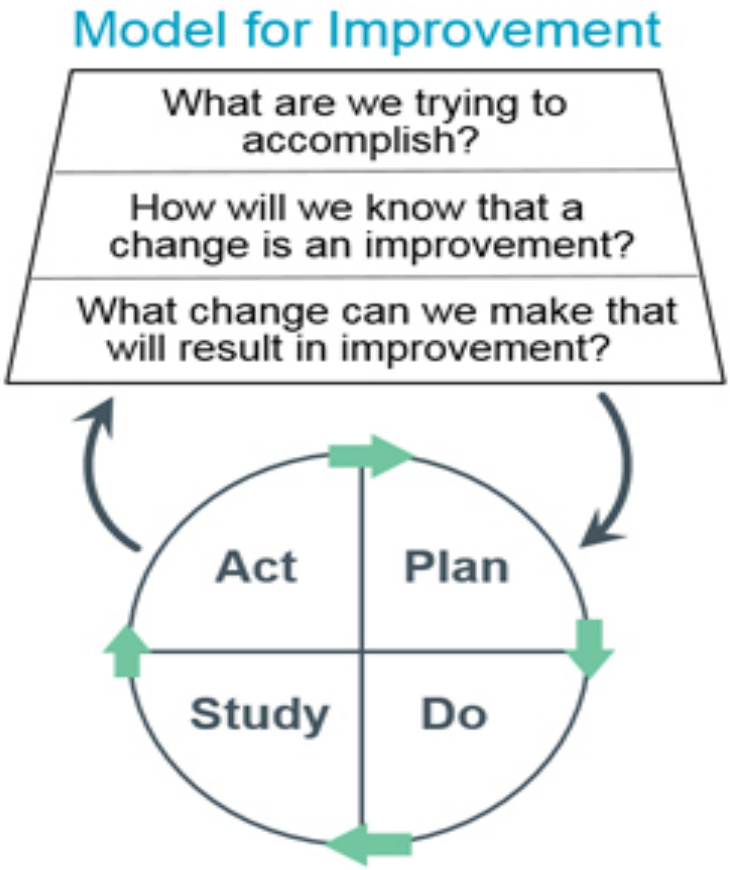
| Author | Title | Synthesis | Findings | Appraisal |
|---|---|---|---|---|
| | | | • Two major revisions to WMI were identified as important: 1st change related to recognition of importance given to administration and communication by the collaborators and added as subtasks in modules; 2nd recombination of modules 2a and 2b with specific sections related to consent, treatment and follow-up incorporated into Module 2 | |

## Appendix B

*List of Central Processes and Special Procedures*

| **Central Processes**<br>OPAL score accounts for the items below | **Special Procedures**<br>One (1) point will be added to the OPAL score for each timepoint requiring the items below |
|---|---|
| • Protocol and Informed Consent Form review<br>• Submission to Institutional Committee(s) for review/approvals: initial review, amendments, continuing reviews, reportable events, exceptions/deviations<br>• Study related training (e.g., SIV, EDC)<br>• Triaging study related patient and provider inquiries<br>• Screening and enrollment of subjects<br>• Scheduling study patient visits<br>• Coordinating study related procedures and visits<br>• ECHO/MUGA<br>• Insertion of apheresis catheter<br>• Leukapheresis<br>• Administration of Investigational Product (IP) – intravenous infusion<br>• Preparation for monitoring visits and audits<br>• Safety monitoring and reporting<br>• Source documentation<br>• Data entry and reconciliation<br>• Electronic medication orders (Beacon) | • Antigen Expression Testing<br>• Bone Marrow Biopsy<br>• Tumor Biopsy<br>• Nuclear Imaging<br>• Neuro Evaluation<br>• Insertion of specialty catheter for intrapleural, intraperitoneal or intratumoral administration of IP<br>• Administration of IP – intrapleural and intraperitoneal infusion or intratumoral injection<br>• Lymphodepleting chemotherapy<br>• Retreatment – planned retreatment of IP after the initial study treatment at Day 0<br>• Fractionated dosing or multiple administration study visits<br>• Multiple treatment cohorts |

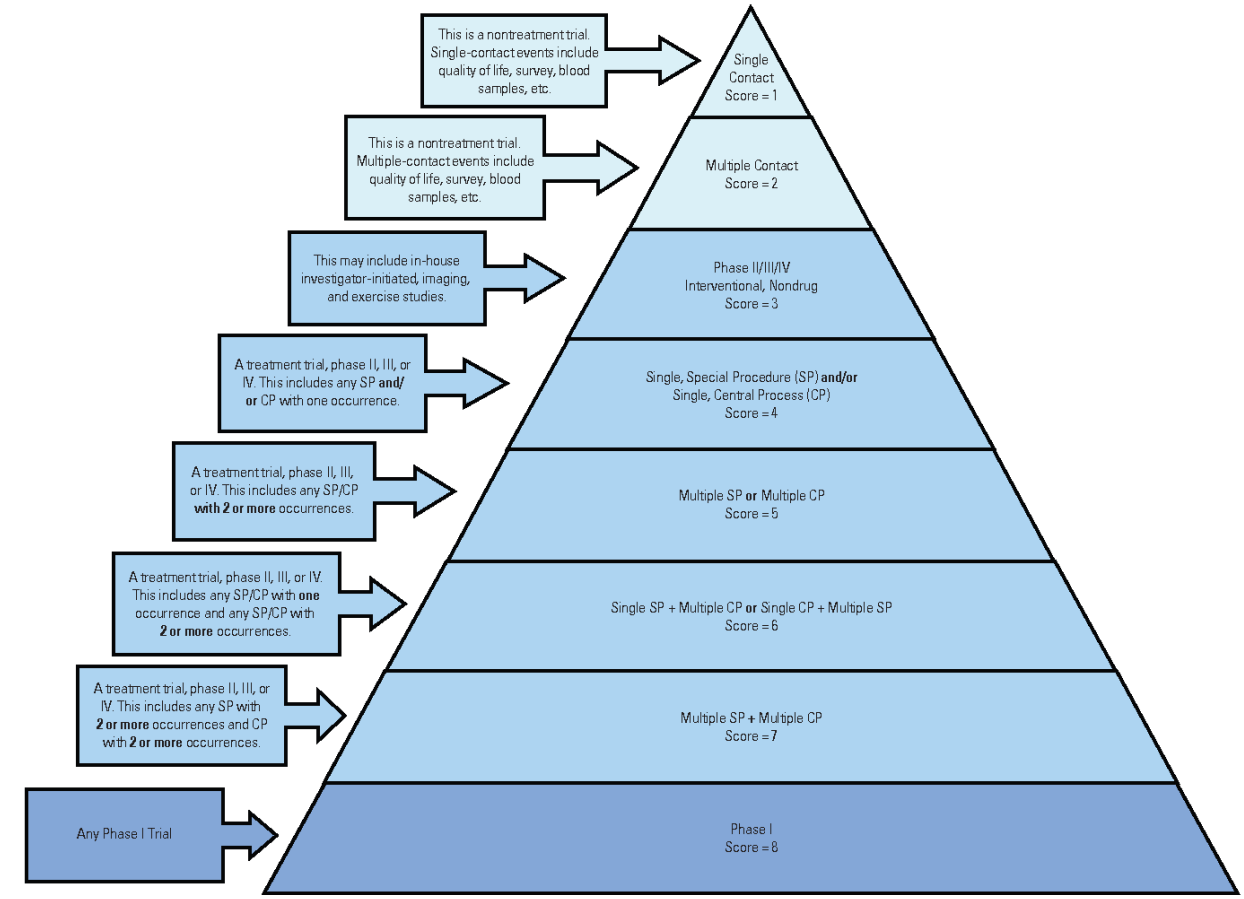
**Appendix C**




*Note.* IHI Model for Improvement. Reprinted with permission from John Wiley and Sons. Copied from the Institute for Healthcare Improvement. Science of improvement: Testing changes. IHI n.d. http://www.ihi.org/resources/Pages/HowtoImprove/ScienceofImprovementTestingChanges.aspx

## Appendix D

*OPAL Pyramid Scale*




*Note.* The Ontario Protocol Assessment Level (OPAL) pyramid scale. Copied from "Ontario Protocol Assessment Level: Clinical Trial Complexity Rating Tool for Workload Planning in Oncology Clinical Trials", by Smuck, B. et al., 2011, *Journal of Oncology Practice, 7*(2), p.82. Copyright 2011 by American Society of Clinical Oncology. Reprinted with permission.

## Appendix E

*Job Stress Questionnaire Dimension 1*

| Job Requirements: Dimension 1 | Rarely | Occasionally | Sometimes | Fairly Often | Very Often |
|---|---|---|---|---|---|
| How often does your job require you to work very fast? | 1 | 4 | 8 | 4 | |
| How often does your job require you to work very hard? | 1 | 1 | 5 | 7 | 2 |
| How often does your job leave you with little time to get things done? | 3 | 3 | 6 | 4 | |
| How of often is there a great deal to be done? | 1 | 1 | 7 | 7 | 1 |
| How often is there marked increase in the workload? | 1 | 3 | 10 | 3 | |
| How often is there marked increase in the amount of concentration required on your job? | | 2 | 6 | 7 | 1 |
| How often is there a marked increase in how fast you have to think? | 1 | 4 | 6 | 5 | |
| How often does your job let you use the skills and knowledge you learned in school? | 3 | 3 | 3 | 5 | 2 |
| How often are you given a chance to do the things you do the best? | 2 | 1 | 4 | 7 | 2 |
| How often can you use the skills from your previous experience and training? | | 3 | 4 | 8 | 1 |

## Appendix F

*Job Stress Questionnaire Dimension 2*

| Workload and Responsibility: Dimension 2 | Hardly Any | A Little | Some | A Lot | A Great Deal |
|---|---|---|---|---|---|
| How much slowdown in the workload do you experience? | 1 | 8 | 7 | | |
| How much time do you have to think and contemplate? | | 1 | 11 | 3 | 1 |
| How much workload do you have? | | | 6 | 10 | |
| What quantity of work do others expect you to do? | | | 9 | 7 | |
| How much time do you have to do all your work? | | 2 | 6 | 8 | |
| How many projects, assignments, or tasks do you have? | | | 6 | 9 | 1 |
| How many lulls between heavy workload periods do you have? | 1 | 5 | 9 | 1 | |
| How much responsibility do you have for the future of others? | 4 | 2 | 8 | 1 | 1 |
| How much responsibility do you have for the job security of others? | 8 | 1 | 5 | 1 | 1 |
| How much responsibility do you have for the morale of others? | 4 | 1 | 7 | 3 | 1 |
| How much responsibility do you have for the welfare and lives of others? | 3 | 3 | 8 | 2 | |

## Appendix G

*Job Stress Questionnaire Dimension 3*

| Mental Demands: Dimension 3 | Strongly Agree | Slightly Agree | Slightly Disagree | Strongly Disagree |
|---|---|---|---|---|
| My job requires a great deal of concentration. | 11 | 4 | | 1 |
| My job requires me to remember many different things. | 12 | 2 | | 1 |
| I must keep my mind on my work at all times. | 6 | 10 | | |
| I can take it easy and still get my work done. | | 2 | 9 | 5 |
| I can let my mind wander and still do the work. | | 1 | 9 | 5 |

## Appendix H

*Job Satisfaction Dimension 4*

| **Job Satisfaction: Dimension 4** | | |
|---|---|---|
| **Knowing what you know now, if you had to decide all over again whether to take the type of job you now have, what would you decide?** | **I would decide without hesitation to take the same job.**<br>**I would have some second thoughts.**<br>**I would decide definitely not to take this type of job.** | **14**<br>**2** |
| **If you were free right now to go into any type of job you wanted, what would your choice be?** | **I would take the same job.**<br>**I would take a different job.**<br>**I would not want to work.** | **11**<br>**5** |
| **If a friend of yours told you he/she was interested in working in a job like yours, what would you tell him/her?** | **I would strongly recommend it.**<br>**I would have doubts about recommending it.**<br>**I would advise against it.** | **15**<br><br>**1** |
| **All in all, how satisfied would you say you are with your job?** | **I am very satisfied.**<br>**I am somewhat satisfied.**<br>**I am not too satisfied.** | **8**<br>**8** |