Intervention Effectiveness Research: Quality Improvement and Program Evaluation

Karen A. Monsen

Springer

Intervention Effectiveness Research: Quality Improvement and Program Evaluation

Karen A. Monsen

Intervention Effectiveness Research: Quality Improvement and Program Evaluation

Springer

Karen A. Monsen
School of Nursing
University of Minnesota System School of Nursing
Minneapolis
Minnesota
USA

ISBN 978-3-319-61245-4 ISBN 978-3-319-61246-1 (eBook)
DOI 10.1007/978-3-319-61246-1

Library of Congress Control Number: 2017953368

© Springer International Publishing AG 2018
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

For Wayne

Foreword

The thirst for discovery, improved outcomes, and quality of life are embedded in the aspirations of individuals, organizations, and whole systems. Whether it is our commitment to person-centered care, organizational efficiency, the learning health system, or planetary health, we recognize that science provides a vigorous path to discovery. Further, impact of the science rests in our undaunted embrace of bold challenges and unwavering connection to the everyday lives of people, families, and communities.

Monsen's "Intervention Effectiveness Research: Quality Improvement and Program Evaluation" is an exemplar in addressing the thirst for discovery with an unwavering linkage to impact on the lives of people, communities, and systems. While she clarifies definitions of intervention effectiveness research, quality improvement activities, and program evaluation at the start, Monsen is clear to address what is at the heart of the matter. She asserts that "the need to know if an efficacious or research-based intervention is effective in a real-world setting, producing the intended and desired outcomes," is this heart. Chapters 1–8 capture the Problem-Intervention-Outcome Meta-Model and the process of discovery form design, tools, data analysis, interpretation, and pattern detection, through ethical considerations. Consistent with her "heart of the matter" is a compilation of Chapters 9–17 that provides a practical guide based on the Problem-Intervention-Outcome Meta-Model. A plethora of support tools and practical recommendations are provided including worksheets and statements, standard methods for implementation, and exemplars for dissemination to broad audiences through reports, presentations, and publications. The work is complemented by a useful guide to abbreviations and excellent referencing.

Most compelling is Monsen's ability to incorporate information and implications of our ongoing scientific revolution within "Intervention Effectiveness Research: Quality Improvement and Program Evaluation." She incorporates big data resources, methods, partnership, team-based science, and other aspects of this revolution at a very time when the human tendency is to retract, contract, and quiet creativity.

Monsen sets the stage for the science of improvement, emphasizing innovation and rapid-cycle testing in health care. She examines the cyclical process of evaluating interventions, changing practice based on that evaluation, and reevaluating again. She notes an ongoing, never-ending process because evidence changes,

individuals respond in diverse ways to interventions in the real world, and healthcare quality must be continually updated.

This book invites the scientist and the implementer to the table of making a difference in real lives and to a path of continuous discovery. And as T. S. Eliot reminds us: “We shall not cease from exploration, and the end of all our exploring will be to arrive where we started and know the place for the first time.”

Minneapolis, MN, USA Connie White Delaney, PhD, RN, FAAN, FACMI

Preface

In 1997, I unwittingly embarked upon a life-changing journey launched by the charge of the director of Washington County Minnesota Department of Public Health, Mary McGlothlin, to "computerize documentation and show outcomes." Now 20 years later, I'm still passionate about that goal, still learning new ways of examining data for important clinical outcomes, and still being inspired by discovery. With every study, new evidence about people, their health outcomes, and their care providers comes to light – important new evidence for caring and improving care. Those of us who mine the data are privileged to see and understand the stories that are hidden in the data. We have the responsibility to make those stories visible and to apply what we learn. Together we can health optimize outcomes based on what these stories tell us.

Over all these years, I have come to see that the work of understanding healthcare interventions and outcomes in practice settings is a specialized sliver of the continuum of research and evaluation that exists in health care from acute care to public health. I have had the privilege of working across health care settings, and have seen that essentially, we're all doing our best to make sure we use resources in the best way to improve health for people and populations. In this book, I distill the worldviews and specialized terminologies used in intervention effectiveness research, quality improvement, and program evaluation into a simple meta-model that can guide efforts to understand the real-world intervention-outcome relationships in the context of population, practice, and program. My hope is that this distillation will simplify and advance the study of interventions and outcomes, and lead to an easier path from data to practice and policy. I also hope that these ideas will spur more inquiry, and that those who leverage the power of data will know the joy of discovery that is sure to be found at the end of each project.

This book draws on my personal experiences, lessons learned, and epiphanies from my career as a nurse, manager, educator, and scientist. I hope to provide readers with easy entrée to data-based discovery. Having learned that structure is useful, I've provided a lot of structure in the worksheets and checklists herein. I've also learned that structure is best with some wiggle-room, and thus the worksheets and checklists are also intended to be general guides that can help readers find what they need in order to succeed. I'm delighted and honored to share insights that I've found to be helpful.

I would like to express my gratitude to my excellent mentors, especially to Dean Connie White Delaney, who kindly provided the foreword to the book; to Madeleine J. Kerr, my excellent advisor and colleague; and to Karen S. Martin, who took an interest in and championed my work with data. The combination of their visions for using data to improve health is the foundation for my inspiration and the reason for this book.

Three individuals read manuscript drafts and provided excellent feedback. It is a much better book than it would have been because they did so. I am grateful to an exceptionally inspiring, brilliant, and generous colleague, Daniel J. Pesut, whose guidance and ideas contributed to the new perspective that will shape future work based on this model. I am deeply indebted to the astute and creative Michelle A. Mathiason, whose insights, suggestions, and perspective gave me confidence and helped me to see the possibilities for this book. I am grateful to Doug Toft who knows books and what it takes to write them; whose advice and encouragement for the authorship journey made finishing this book possible.

Finally, I wish to thank my dear husband and family, and wonderful friends, students, and colleagues who provided love, wisdom, and inspiration. You gave me the courage to write these words. As the sign on my office wall says, “Words are the voice of the heart.” The words in this book are for you. From the bottom of my heart, I thank you.

Minneapolis, MN, USA Karen A. Monsen

Contents

List of Abbreviations

CINAHL	Cumulative Index to Nursing and Allied Health Literature
CSV	Comma-separated values
EDA	Exploratory data analysis
EHR	Electronic health record
IHI	Institute for Healthcare Improvement
NLM	National Library of Medicine
NMWC	Nurse-managed wellness center
PDSA	Plan-do-study-act
PICOT format	P stands for population, I stands for intervention, C stands for comparison, O stands for outcome, and T represents time
PIO MM	Problem-Intervention-Outcome Meta-Model
PHN	Public health nurse
P_{TimeX}	Measure of a problem (P) at a given time (X)
QI	Quality improvement
SQUIRE	Standards for Quality Improvement Reporting Excellence

Part I

Introduction to Intervention Effectiveness Research, Quality Improvement, and Program Evaluation

Key Concepts, Definitions, and Frameworks

1

1.1 Introduction

This book is intended to be a practical guide designed for those who are embarking on a journey of demonstrating care quality and outcomes the at the intersection of research and practice, alternatively described as intervention effectiveness research, quality improvement activities, and program evaluation. We acknowledge that in this space are dynamically interacting factors that affect all health outcomes at the individual, family, community, and system levels. These factors range from the genome to the political and include physical and mental characteristics as well as social processes within and external to the individual and the health system [1].

Efforts to understand the effectiveness of interventions or programs in relationship to social and behavioral factors will benefit from taking a high-level look at problems in the broad context of population and setting, and taking into account the diversity of intervention factors known to be efficacious for the population [1]. This perspective is conceptualized in the Problem-Intervention-Outcome Meta-Model (PIO MM) described in Chapter 2. Transcending the differences among various worldviews of research, quality improvement, and evaluation allows us to think more clearly about the fundamental concepts and realize our potential to study and act to improve population health [1].

We are living in an era launched by the advent of technology and computing capacity in which intervention effectiveness research, quality improvement activities, and program evaluation may rapidly advance knowledge discovery in health care by leverage existing data across all levels, especially data from electronic health records (EHRs) [2–4]. Such data enable us to identify important novel patterns and conduct comprehensive evaluations of clinical outcomes. We are beginning to tap into the potential to generate useful information, knowledge, and practice wisdom [2, 5–10].

Projects that leverage existing data may be considered within the recently described *fourth paradigm of research* [9], and may include or transcend classical

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_1

theoretical and empirical research paradigms, such as randomized controlled trials [9]. Designs are variations on the single-group before and after design and may incorporate quantitative, qualitative, and big data methodologies, described in Chapter 3. While we do not discourage use of prospective data collection, the methods we describe in this book are amenable to use with existing datasets.

1.2 Definitions and Descriptions of Intervention Effectiveness Research, Quality Improvement, and Program Evaluation: What They Have in Common and How They Differ

Intervention effectiveness research is a formal research step on a continuum between bench science and translation and use of an intervention in practice (bench discovery—safety and efficacy in controlled setting—real world effectiveness—translation to practice) [11]. It is intended to test *safe* and *efficacious* interventions in real-world or nearly real-world settings to understand whether the intervention works in less controlled settings—or to test *effective* interventions in additional real-world settings in which effectiveness has not been demonstrated. It is likely to rely on a validated theoretical framework. It focuses on generalizability, because it is already known that the intervention a good and safe thing to do in a controlled setting. In such cases external validity is as important as internal validity [1]. Positive findings in intervention effectiveness research lay the foundation for translation of the intervention into practice, and may measure the value of providing the intervention. Intervention Effectiveness Research answers the question: Was this intervention associated with an expected improvement in a health problem in a real world setting? [11].

1.2.1 What Is Quality Improvement?

Quality improvement is defined as actions that lead to measurable improvement in health care services and the health status of targeted groups [12]. Health care quality improvement is a special case of intervention effectiveness evaluation that seeks to rapidly implement and test interventions that work to improve practice using readily available clinical data [12–14]. Quality improvement addresses health services inefficiencies and tests individual and health system responses to changes in processes that are implemented based on best evidence [14]. The Institute of Medicine (IOM) defines quality in health care as *a direct correlation between the level of improved health services and the desired health outcomes of individuals and populations* [12]. Typically there is a feedback loop from results to the next evaluation in an ongoing, iterative manner. Quality Improvement evaluation answers the question: Was this intervention/process change associated with an improvement in outcomes in this hospital or health system? [12–14].

1.2.2 What Is Program Evaluation?

Program evaluation is a structured process intended to measure whether or not an intervention or program was successful in meeting its goals [15–17]. This happens routinely in public health and clinical practice, as clinical leaders or program planners consult stakeholder; set program goals; develop program evaluation plans; collect, analyze, and interpret data; and share lessons learned [18]. Often these efforts are used in government, non-profit, and other real-world settings to examine the outcomes of evidence-based programs, and are paired with process evaluation which measures the way the intervention (or program) is provided [15–17]. In some cases, program evaluation further extends outcomes from the immediate changes related to the intervention, to impacts, or sustained and extended changes related to the interventions. Positive findings in program evaluation may be useful to support funding, policy, and practice change. Program evaluation answers the question: Was this intervention/program associated with improvement in individual or population health? [15–16].

1.3 How Intervention Effectiveness Research, Quality Improvement Activities, and Program Evaluation Are Similar

These approaches are similar in in that they focus on real world settings, measure interventions and outcomes over time, and have practice and policy implications. Despite arising from very different traditions, intervention effectiveness research, quality improvement activities, and program evaluation include two fundamentally equivalent components: implementing an evidence-based intervention in the real world, and measuring outcomes associated with that intervention.

The three approaches are also similar in that they do not or cannot implement controls that typically provide confidence in the findings of a typical clinical trial. The reason for this in intervention effectiveness research is that we need to know if efficacious interventions (proven in controlled settings) are also effective (proven in uncontrolled settings)—and thus may be generalizable beyond the research environment into practice. The reason for this in practice settings (quality improvement and outcomes evaluation) is that it is not possible or ethical to implement research controls such as randomization to control groups or treatment groups in clinical settings in which people receive the best known care in response to each individual's healthcare needs. There is no good rationale for creating 'no care' or 'sub-optimal care' comparison groups in a prospective intervention effectiveness, quality improvement, or program evaluation project given the knowledge that the intervention is effective or efficacious.

Another way in which the approaches are similar is the potential for re-use of clinical and public health data related to the intervention in order to evaluate outcomes. Such data are embedded within existing information systems and structures and may be accessed by researchers, quality improvement leaders, and evaluators

using or establishing secure data transfer and storage mechanisms. It is critical to cultivate and maintain positive working relationships between those who will conduct the studies and those who will provide the data so that the process is supported; the outcome analysis is informed and validated in practice; and the findings are shared where they are most relevant.

1.4 How Intervention Effectiveness Research, Quality Improvement, and Outcome Evaluation Are Different

Differences between these three perspectives are mainly in the phrases used by the cultures from which they have grown up [11–16]. There are several interpretations of each perspective; and scholarship has evolved over time in various settings, influencing the ways they have been described. Recognizing the underlying concepts and how they are described helps to discriminate between the projects and studies that are and are not reflective of research and evaluation intended to demonstrate care quality and outcomes. In this section we consider approaches that differ from intervention effectiveness research, quality improvement activities, and program evaluation.

1.4.1 Translational Research

Some scholars describe the continuum of translational research in two stages (T1, T2) from laboratory-to-humans, in which T1 includes research related to the development of the intervention, and T2 includes quality improvement and evaluating the impact of intervention [19]. Others describe translational research in four stages (T1, T2, T3, T4) in which the T3 stage research yields knowledge about how interventions work in real-world settings [20]. Thus, if describing intervention effectiveness research in terms of translational research, it is necessary to provide the overall context of the defined stages used by the organization.

1.4.2 Quality Improvement (QI)

Quality improvement processes have been variously described by terms such as total quality management, process improvement, and continuous quality improvement [13, 21–22]. Quality improvement studies may be the responsibility of organizational departments in health care systems, and the approaches used within these systems may differ. Many quality improvement studies focus on improving one or more aspects of the Institute for Healthcare Improvement "Triple Aim" defined as improving the individual experience of care (including quality and satisfaction), improving the health of populations, and reducing the per capita cost of health care

[22]. Quality improvement studies may be developed for one or more aspects of the Triple Aim, particularly improving the health of populations may include quality improvement studies that focus on improving health outcomes [23].

1.4.3 Six Sigma Quality Improvement

Like the PDSA cycle, Six Sigma process consists of steps to define, measure, analyze, and improve healthcare quality [24–27]. It is a data-driven process derived from the manufacturing industry that has been shown to reduce errors and improve timeliness of intervention processes at individual and health system levels:

- Define: Who do we serve, and what do they want? What are the capabilities of the process being measured? What are our objectives?
- Measure: What will improvement look like? On what data will our efforts be measured? Metrics used to define performance in a healthcare organization often include service level, service cost, individual satisfaction, and clinical excellence.
- Analyze: Collect data and analyze using proven tools.
- Improve: Implement modifications to improve the process.
- Control: Monitor performance to maintain improvement.

The Six Sigma approach to quality improvement is a proprietary method that involves a hierarchy of training levels across an organization [24–25]. Quality improvement studies may be developed using the Six Sigma methodology with particular attention to the evidence for implementing changes in the interventions that will be implemented in the process being studied. It should be noted that the Six Sigma approach may be limited by its purposeful focus on reducing variability; which is useful in manufacturing but may be less so in healthcare [26].

1.4.4 Health Services Research

Health services research is defined as a "multidisciplinary field of scientific investigation that studies how social factors, financing systems, organizational structures and processes, health technologies, and personal behaviors affect access to health care, the quality and cost of health care, and ultimately, our health and well-being." [28] It spans disciplines and health system levels in generating new knowledge to make health care affordable, safe, effective, equitable, accessible, and consumer-centered. For example, results of health services findings may enable providers and the people they serve to make better decisions, and/or design health care benefits and inform policy [28]. Applied health services research may include studies that examine intervention effectiveness and quality.

1.4.5 Big Data in Health Care Research

Big data is defined by high levels of volume, velocity, and variety of data sources or types that make analysis difficult using standard statistical methods and computing techniques [29–30]. Newly abundant data streaming from consumer devices and generated through real-time documentation in EHRs are becoming available and have potential for use in intervention effectiveness research and quality improvement studies. The goal of big data research may be to provide real-time decision support for consumers or clinicians; or to identify patterns that can be evaluated to generate new knowledge to improve healthcare quality and outcomes as in the Triple Aim [22, 29]. Challenges in using big data are many and include issues of data structure, security, data standardization, storage and transfers, and managerial skills such as data governance [30]. Data science, a specialized field that is rapidly emerging in response to these challenges, contributes valuable expertise for data analytics and is becoming more available across sectors including healthcare. Healthcare research teams that desire to use big data sources and methods should include individuals who are skilled in big data management and analysis techniques as well as clinical experts [30].

1.4.6 Program Evaluation

Diverse types of evaluation approaches measure different aspects of a program designed to promote health, and/or prevent or mitigate disease [17, 31–32]. By definition, programs provide evidence-based or best practice interventions to individuals or communities that may benefit in relationship to an identified health need or risk of disease. As such, the World Health Organization noted that the "use of randomized controlled trials to evaluate health promotion initiatives is, in most cases, inappropriate, misleading and unnecessarily expensive" [33]. Thus the science of program evaluation has expanded over the past decades. Aspects of these programs that are commonly evaluated include planning (formative evaluation), implementation process and quality (process evaluation), short term results (program evaluation), and long term or system/policy results (impact evaluation) [17, 31–32].

- Formative evaluation: Formative evaluation occurs during program development and implementation. It provides information on achieving program goals or improving your program.
- Process evaluation: Process evaluation is a type of formative evaluation that assesses the type, quantity, and quality of program activities or services.
- Program evaluation: Program evaluation can focus on short- and long-term program objectives. Appropriate measures demonstrate changes in health conditions, quality of life, and behaviors. Also called Summative evaluation.
- Impact evaluation: Impact evaluation assesses a program's effect on participants. Appropriate measures include changes in awareness, knowledge, attitudes, behaviors, and/or skills [13].

Rigor in program evaluation is essential to ensure confidence in the findings. Rigorous evaluation may be ensured by implementation (with fidelity) of an intervention approach for which effectiveness has been established through previous research, and conducting evaluation according to established evaluation procedures [17].

1.4.7 Implementation Research

Implementation research is the scientific inquiry into questions concerning implementation—the act of carrying an intention into effect, which in health research can be policies, programs, or individual practices (collectively called interventions) [34]. Implementation outcome variables describe the intentional actions to deliver services and are measures of intervention acceptability, adoption, appropriateness, feasibility, fidelity, implementation cost, coverage, and sustainability [35]. As part of an intervention effectiveness research, quality improvement activities, or program evaluation project, implementation research is a way of understanding how well an implementation was implemented, and results may be associated with other study or program findings. We include these variables as process variables related to our intervention effectiveness research, quality improvement activities, and program evaluation—knowing that these may be contributing factors to the overall outcomes [34].

1.5 Definitions of Similar Sounding Terms and What This Book Does Not Attempt

1.5.1 Comparative Effectiveness Research

The Institute of Medicine defined comparative effectiveness research as "Comparative effectiveness research is the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor or improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers, and policy makers to make informed decisions that will improve health care at both the individual and population levels." [36] The key reason we are not focusing on CER is that comparing two different interventions goes beyond the simple evaluation of the effectiveness of a single evidence-based intervention or program.

1.5.2 Implementation Science Research

Implementation science is the study of methods to promote the integration of research findings and evidence into healthcare policy and practice [37]. The types of studies in this newly emerging field may vary may vary according setting and sponsor, though it is generally understood that the intent of implementation science is to investigate and address major social, behavioral, economic, management challenges

that are *barriers to effective implementation*, as well as testing *new* approaches to improve health programming, as well as discover *causal relationships* between interventions and outcomes [37]. The key reasons we are not focusing on implementation science research is that researching the *methods of studying intervention implementation* are not the focus of intervention effectiveness research, quality improvement activities, and program evaluation; likewise given our focus on observational studies, it is not possible to determine a causal relationship between interventions and outcomes.

1.5.3 Dissemination Science

Dissemination science is the study of the targeted distribution of information and intervention materials to a specific public health or clinical practice audience [38]. It is a corollary of Implementation Science that aims to examine and promote changes at the intersection of health services, delivery systems, and *communities*; investigating the reach, relevance, uptake, and diffusion of interventions to influence population health [38]. Beyond the scope of intervention effectiveness research, quality improvement activities, and program evaluation, dissemination science addresses the need for research testing *the approaches to scaling up and sustaining effective interventions* [38]. Thus, dissemination science necessarily builds on the body of evidence generated by intervention effectiveness research, quality improvement activities, and program evaluation and aligns with stage T4 in the continuum of translational research.

1.6 Frameworks to Support Intervention Effectiveness Research, Quality Improvement Activities, and Program Evaluation

Intervention effectiveness research, quality improvement activities, and program evaluation depend on various theories, models, or frameworks such as Logic Models, Conceptual Frameworks, or Theoretical Frameworks, to describe the way in which intervention(s) relate to outcome(s) for a given population or context [11, 15–16, 39–42].

1.6.1 Theory

Theory is defined as "a set of statements or principles devised to explain a group of facts or phenomena, especially one that has been repeatedly tested or is widely accepted and can be used to make predictions about natural phenomena." [43] A theory may be derived inductively or deductively, but is not discovered. Theory may be used to guide research, and findings of research based on a given theory may

be compared to increase generalizability across studies. The advancement of the theoretical discourse of the discipline is furthered through application and reflection on the use of theory in diverse studies.

Intervention effectiveness has been studied in relationship to theory, and it has been found that interventions with a theoretical basis may be more effective compared to those without, and that interventions guided by multiple theories may be even more effective [44]. Health and social sciences have proposed numerous theories that have been used to greater and lesser extents in scientific inquiry and practice [44].

Improved healthcare outcomes depend on changes in clinician and/or consumer behavior in the context of real life, and thus theories that explain how interventions affect behavior change that incorporate social aspects are particularly useful. The most common social and behavioral change theories used in health promotion and public health are: Social Cognitive Theory; Transtheoretical Model/Stages of Change, Health Belief Model, and the Theory of Planned Behavior [44–48]:

- Social Cognitive theory explains human behavior in terms of three-way, dynamic, reciprocal interaction that relates personal factors, environmental influences, and behavior in a nested model [45]
- Transtheoretical model/stages of change describes stages of precontemplation, contemplation, determination, action, relapse, and maintenance [46]
- Health belief model explains perceptions of people in regard to use of health services. Core constructs include perceived susceptibility, severity, benefits, and barriers; cues to action, and self-efficacy [47]
- The theory of planned behavior links beliefs and behavior [48].

Nursing theories intended to explain how nursing interventions may affect outcomes and/or used to guide and inform interventions are numerous. Some commonly used nursing theories are the Health Promotion Model, Health Behavior Theory, Behavior System Model of Nursing, and Theory of Planned Action [49–52]

- Health Promotion Model relates multidimensional individual factors and behavioral factors interacting within the environment to behavior outcomes and well-being [49].
- Integrated Theory of Health Behavior Change relates knowledge and beliefs, increasing self-regulation skills and abilities, and social facilitation with self-management behaviors and in turn, improved health status [50].
- Behavior System Model of Nursing incorporates biological and behavioral systems within society taking into account motivation and goals [51].
- The theory of reasoned action and theory of planned behavior focus on individual motivational factors that relate to behaviors, such as intention, and incorporate constructs such as perceived control [52].

Use of one or more theories may be combined with other concepts and ideas in a logic model, conceptual framework, or theoretical framework to provide an overview of and model for a proposed project.

1.6.2 Logic Models

A logic model is defined as a graphical and/or textual representation of how a project or program is intended to work. It may include program elements and measures, as well as linkages of processes to outcomes and the theoretical assumptions that underlie the program [15, 39]. It has been shown to be useful for program planning in public health as well as primary care and other settings [15, 39]. Logic models may be used to describe resources, activities and anticipated results for each project goal. A well-defined logic model explicates assumptions, measures, and processes that aid in planning and implementation of the evaluation project [15, 39].

1.6.3 Theoretical Framework

A theoretical framework is defined as a group of related ideas that provides guidance to a research project [40]. The theoretical framework for a given project must illuminate the theories and concepts that are relevant to the intervention and desired outcomes for the population of interest, within the specific context. Thus the theoretical framework is not found in the literature; rather, aspects of the theoretical literature may be assembled into a coherent whole for purposes of guiding study planning and implementation [53].

1.6.4 Conceptual Framework

A conceptual framework is defined as system of concepts, assumptions, expectations, beliefs, and theories that supports and informs research [54] Thus, the conceptual framework extends beyond theoretical discourse and explores the underlying thought processes that may bias the project. By creating such a systematic approach to examining these aspects of thought, conceptual frameworks clarify and integrate philosophical, methodological and pragmatic aspects of the study [55].

In Chapter 2, we will further extend the notion of theory for intervention effectiveness research, quality improvement activities, and program evaluation as we examine a meta-model that describes fundamental problem-intervention-outcome concepts and relationships.

Reflection Questions

- In your own words, explain the differences between intervention effectiveness research, quality improvement activities, and program evaluation.
- Given the context of your work, which perspective is most comfortable to you and your colleagues?
- Are distinctions between the approaches necessary, and if so, why?

References

1. Smedley BD, Syme SL (2001) Promoting health: intervention strategies from social and behavioral research. Am J Health Promot 15(3):149–166
2. Russo E, Sittig DF, Murphy DR, Singh H (2016) Challenges in patient safety improvement research in the era of electronic health records. Healthc (Amst) 4(4):285–290
3. Hersh WR, Cimino J, Payne PR, Embi P, Logan J, Weiner M, Bernstam EV, Lehmann H, Hripcsak G, Hartzog T, Saltz J (2013) Recommendations for the use of operational electronic health record data in comparative effectiveness research. EGEMS 1(1)
4. Monsen KA, Swenson SM, Klotzbach L, Mathiason MA, Johnson KE (2017) Empirical evaluation of change in public health nursing practice after implementation of an evidence-based family home visiting guideline. Kontakt. doi:10.1016/j.kontakt.2017.03.002
5. Ronquillo C, Currie LM, Rodney P (2016) The evolution of data-information-knowledge-wisdom in nursing informatics. Adv Nurs Sci 39(1):E1–E8
6. Olsen L, McGinnis JM (2010) Redesigning the clinical effectiveness research paradigm: innovation and practice-based approaches: workshop summary. National Academies Press, Washington, DC
7. Jensen PB, Jensen LJ, Brunak S (2012) Mining electronic health records: towards better research applications and clinical care. Nat Rev Genet 13(6):395–405
8. Mayer-Schönberger V, Cukier K (2013) Big data – a revolution that will transform how we live, work and think. John Murray Publishers, London
9. Hey T, Tansley S, Tolle K (2009) The fourth paradigm: data intensive scientific discovery. Microsoft Research, Redmond, WA
10. Schneeweiss S (2014) Learning from big health care data. N Engl J Med 370(23):2161–2163
11. Drolet BC, Lorenzi NM (2011) Translational research: understanding the continuum from bench to bedside. Transl Res 157(1):1–5
12. Health Resources and Services Administration. Quality improvement [Internet]. [cited 2017 May 12]. Available from: https://www.hrsa.gov/quality/toolbox/methodology/qualityimprovement/
13. Institute for Healthcare Improvement. The science of improvement: how to improve [Internet]. [cited 2017 May 12]. Available from: http://www.ihi.org/resources/pages/howtoimprove/scienceofimprovementhowtoimprove.aspx
14. Richardson WC, Berwick DM, Bisgard JC, Bristow LR, Buck CR, Cassel CK (2001) Crossing the quality chasm: a new health system for the 21st century. National Academies Press, Washington, DC
15. Silverman B, Mai C, Boulet S, O'Leary L Logic models for planning and evaluation. The Division of Birth Defects and Developmental Disabilities (DBDDD), Centers for Disease Control and Prevention (CDC), Atlanta, GA

16. Krueger RA. Evaluation [Internet]. [cited 2017 May 12]. Available from: http://www.tc.umn.edu/~rkrueger/evaluation.html
17. Brownson RC, Baker EA, Leet TL, Gillespie KN, True WR (2010) Evidence-based public health. Oxford University Press
18. Centers for Disease Control and Prevention [US] (1999) Framework for program evaluation in public health. MMWR Recomm Rep 48(RR11):1–40
19. Hawk ET, Matrisian LM, Nelson WG, Dorfman GS, Stevens L, Kwok J, Viner J, Hautala J, Grad O (2008) The translational research working group developmental pathways: introduction and overview. Clin Cancer Res 14(18):5664–5671
20. Khoury MJ, Gwinn M, Yoon PW, Dowling N, Moore CA, Bradley L (2007) The continuum of translation research in genomic medicine: how can we accelerate the appropriate integration of human genome discoveries into health care and disease prevention? Genet Med 9(10):665–674
21. Institute for Healthcare Improvement. IHI Open School. PDSA in everyday life [Internet]. [cited 2017 May 12]. Available from: http://www.ihi.org/Documents/OpenSchoolCourseTranscripts/QI102-Lesson5-Transcript 1.html
22. Institute for Healthcare Improvement. The Institute for Healthcare Improvement triple aim [Internet]. [cited 2017 May 12]. Available from: http://www.ihi.org/Engage/Initiatives/TripleAim/Pages/default.aspx
23. Connor BT. Differentiating research, evidence-based practice, and quality improvement. American Nurse Today 2014 [Internet]. [cited 2017 May 12]. Available from: https://www.americannursetoday.com/differentiating-research-evidence-based-practice-and-quality-improvement/
24. A look at Six Sigma's increasing role in improving healthcare [Internet]. [cited 2017 May 12]. Available from: https://www.villanovau.com/resources/six-sigma/six-sigma-improving-healthcare/#.WPIJ4vnythE
25. What is Six Sigma? [Internet]. [cited 2017 May 12]. Available from: https://www.villanovau.com/resources/six-sigma/what-is-six-sigma/#.WPIWKfnythE
26. NHS Institute for Innovation and Improvement. The handbook of quality and service improvement tools [Internet]. [cited 2017 May 12]. Available from: http://www.miltonkeynesccg.nhs.uk/resources/uploads/files/NHS%20III%20Handbook%20serviceimprove.pdf
27. New to lean Six Sigma [Internet]. [cited 2017 May 12]. Available from: https://www.isixsigma.com/new-lean-six-sigma/
28. Health services research: a definition [US DHHS] [Internet]. [cited 2017 May 12]. Available from: https://www.ahrq.gov/funding/training-grants/hsrguide/hsrguide.html
29. Children's Hospital of Philadelphia Institutional Review Board. Quality improvement vs. research [Internet]. [cited 2017 May 12]. Available from: https://irb.research.chop.edu/quality-improvement-vs-research
30. Raghupathi W, Raghupathi V (2014) Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2(1):3
31. Rural Health Information Hub. Types of evaluation in health promotion and disease prevention programs [Internet]. [cited 2017 May 12]. Available from: https://www.ruralhealthinfo.org/community-health/health-promotion/4/types-of-evaluation
32. Howe P, Degeling D, Hall J (1990) Evaluating health promotion. Maciennan and Petty, Sydney
33. World Health Organization (1998) Health promotion evaluation -recommendations to policymakers. Report of the WHO European working group on health promotion evaluation. WHO, Copenhagen
34. Peters DH, Adam T, Alonge O, Agyepong IA, Tran N (2013) Implementation research: what it is and how to do it. BMJ 347:f6753
35. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A et al (2010) Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 38:65–76
36. Institute of Medicine [US] (2009) Initial national priorities for comparative effectiveness research. The National Academies Press, Washington, DC

37. NIH implementation science information and resources [Internet]. [cited 2017 May 12]. Available from: https://www.fic.nih.gov/researchtopics/pages/implementationscience.aspx
38. Glasgow RE, Vinson C, Chambers D, Khoury MJ, Kaplan RM, Hunter C (2012) National Institutes of Health approaches to dissemination and implementation science: current and future directions. Am J Public Health 102(7):1274–1281
39. Hayes H, Parchman ML, Howard R (2011) A logic model framework for evaluation and planning in a primary care practice-based research network (PBRN). J Am Board Fam Med 24(5):576–582
40. Business dictionary. Theoretical framework. [Internet]. [cited 2017 May 12] Available from: http://www.businessdictionary.com/definition/theoretical-framework.html
41. Yourdictionary.com. Theory dictionary definition|theory defined [Internet]. [cited 2017 May 12]. Available from: http://www.nhlbi.nih.gov/health/resources/heart/heart-cholesterol-hbc-what-html
42. Robson C, McCartan K (2016) Real world research, 4th edn. Wiley, Chichester
43. The free dictionary. Definition of theory [Internet]. [cited 2017 May 12]. Available from: http://www.thefreedictionary.com/theory
44. e-Source Behavioral & Social Sciences Research [Internet]. [cited 2017 May 12]. Available from: Obssr.od.nih.gov
45. Bandura A (2001) Social cognitive theory: an agentic perspective. Annu Rev Psychol 52(1):1–26
46. Prochaska JO, DiClemente CC (1982) Transtheoretical therapy: toward a more integrative model of change. Psychother Theory Res Pract 19(3):276
47. Skinner CS, Tiro J, Champion VL (2015) The health belief model. In: Health behavior: theory, research, and practice. Wiley, San Francisco, pp 75–94
48. Ajzen I (2011) Theory of planned behavior. Handb Theor Soc Psychol 1(2011):438
49. Pender NJ (2011) Heath promotion model manual. University of Michigan
50. Ryan P (2009) Integrated theory of health behavior change: background and intervention development. Clin Nurse Spec 23(3):161
51. Johnson DE (1990) The behavioral system model for nursing. In: Parker ME (ed) Nursing theories in practice. National League for Nursing, New York, pp 23–32
52. Montano DE, Kasprzyk D (2015) Theory of reasoned action, theory of planned behavior, and the integrated behavioral model. In: Health behavior: theory, research and practice, vol 1
53. Larabee RV. Organizing your social sciences research paper. Theoretical framework [Internet]. [cited 2017 May 12]. Available from: 2009 libguides.usc.edu
54. Miles MB, Huberman AM (1994) Qualitative data analysis: an expanded sourcebook. Sage
55. Durham W, Sykes C, Piper S, Stokes P (2015) Conceptual frameworks and terminology in doctoral nursing research. Nurse Res 23(2):8–12

2 Problem-Intervention-Outcome Meta-Model (PIO MM): A Conceptual Meta Model for Intervention Effectiveness Research, Quality Improvement Activities, and Program Evaluation

2.1 Introduction to the Problem-Intervention-Outcome Meta-Model (PIO MM)

In healthcare, we are asked to substantiate the outcomes of our interventions. We are asked: Do we make a difference? And What do we do that makes a difference? In most situations, as described in Chapter 1, we are using an intervention or program in a real-world setting in which we are not able to impose controls or randomize participants. Therefore we must approach the questions from the perspective that there is an expected outcome due to this intervention and we are demonstrating the extent to which the change did (or did not) occur. We need to start by defining the expected difference and the interventions that are known to be efficacious or effective in addressing the problem to achieve the desired outcome.

Modeling change in outcome is complicated by the complex situations and social issues that influence the wellbeing of a population [1–6]. Furthermore, scholars note concerns about poorly described interventions such as health education or social support, that without additional details provide insufficient information about the intervention for replication of the study or meta-analysis of the findings [1]. The Society of Prevention Research (SPR) endorsed a set of standards for intervention effectiveness research that emphasized the clarity of intervention description at a level that would allow others to implement/replicate it [3].

Interventions targeting health behavior change often focusing first on individual factors, such as increasing knowledge, motivation, or skills related to health behavior change [4]. Evidence indicates that behavioral and social interventions can directly impact physiological functioning, and do not merely correlate with positive health outcomes due to improvements in health behavior or knowledge [4]. All such interventions are likely to be more successful when applied in coordinated fashion across multiple levels of influence (i.e., at the individual level; within families and social support networks; within hospitals, schools, work sites, churches, and other community settings; and at broader public policy levels) [4]. Models of intervention must consider individual behavior in a broader social context, with greater attention

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_2

to the social construction of gender, race, and ethnicity, and to ways in which social and economic inequities result in health risks [4].

Thus, as described in Chapter 1, intervention effectiveness research, quality improvement activities, and program evaluation share many common characteristics and have a unique place in health care research because they are embedded within real world settings, making it difficult to impose controls that are the hallmark of clinical trials. Quality Improvement in health care is a special case that utilizes multiple feedback loops in continuous program evaluation. This chapter introduces the fundamental conceptual basis for intervention effectiveness research, quality improvement activities, and program evaluation: the single group before and after design depicted and described in the Problem-Intervention-Outcome Meta-Model (PIO MM).

This section introduces the PIO MM as the meta-model used throughout the book (Fig. 2.1). The PIO MM depicts the notion that an identified sample or population has one or more health problems (P) that may be addressed using one or more interventions (I) to achieve a health outcome (O). The Outcome consists of positive change or benchmark attained in the measure(s) of identified problem(s) over time ($P_{Time2} - P_{Time1}$).

The designs of most intervention effectiveness research, quality improvement activities, and program evaluation studies can be described by variations on the PIO MM. Using the PIO MM will provide a structure for analyzing the study problem and planning the study, enabling identification of key data needs and also gaps that must be addressed to complete the study or evaluation [3].

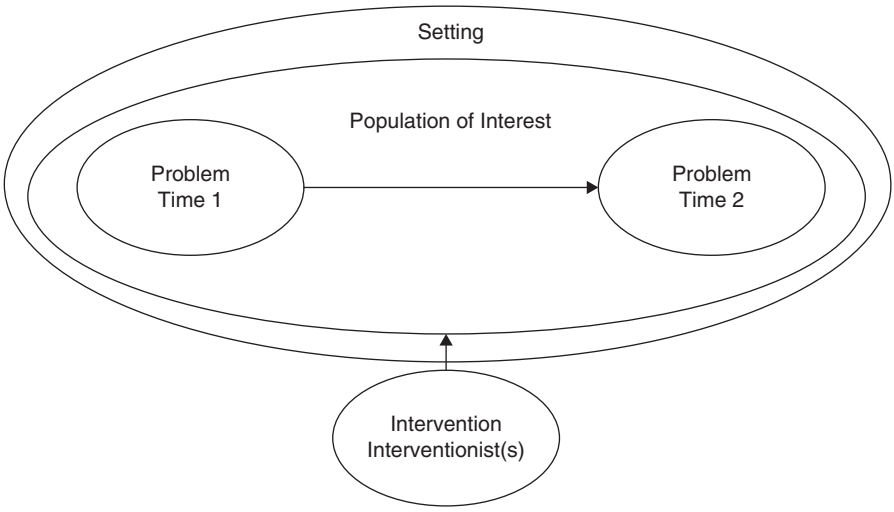

Fig. 2.1 The Problem-Intervention-Outcome Meta-Model depicts the concepts of problem, intervention, population of interest, setting, and interventionist arranged in a relational diagram. Copyright Monsen, K. A. (2016). Used with permission

The PIO MM depicts how theoretically relevant concepts are nested or linked and provides a basis for operationalization of the concepts as variables, and the analysis of data. It may be used to describe how a problem of the population of interest changes in relationship to intervention and contextual factors. This is critical for examining main outcomes as well as complex research problems such as identifying moderating or irrelevant factors, or explaining unexpected data [4].

We will examine definitions of the PIO MM from the perspectives of intervention effectiveness research, quality improvement activities, and program evaluation.

Concept definitions (Health systems research approach):

- Problem—unique concerns, needs, strengths, issues, foci, or conditions that affect any aspect of wellbeing; nursing diagnoses stated from the individual consumer's perspective. It may be broadly conceptualized as a set of neutral concepts that make up a comprehensive, holistic view of individual or community health and wellbeing [7].
- Population of interest—a particular section, group, or type of people; individual, family, or community that receives health care and social services. Synonyms include resident, patient, customer, consumer, and constituency [7].
- Intervention—actions or activities implemented to address a specific problem and to improve, maintain, or restore health or prevent illness [7].
- Outcome—changes or goal attainment in problem measure(s) over time associated with an intervention [7].
- Setting (context)—attributes of the environment or situation that may influence the intervention or outcome [8].
- Time—period or interval over which intervention(s) occur and outcome(s) are measured [9].

2.2 PIO MM and the CDC Logic Model

The PIO MM concepts may be presented in the form of logic model components, the definitions of which align with the public health approach, often used to evaluate publicly funded programs intended to improve health [10–11]. In addition, logic models may include the concepts of Inputs and Outputs, as well as discretized Outcome levels, defined as follows:

- Inputs—resources that go into the program. These are the resources available for a program, such as funding, staff, and leadership, expertise, program infrastructure, scientific knowledge and evidence-based strategies, and partnerships [10–11].
- Outputs—the products or direct services resulting from the program activities. Outputs are the direct evidence of implemented activities; Actions, activities, products, curriculum, etc. of the organization [10–11].
- Short-term outcomes—participant changes in problem-related knowledge, awareness, skills, and status (immediate effects) represent the most immediate

effects attributable to a program, such as changes in learning, knowledge, and attitudes. Short term: These are changes in learning: Knowledge, Opinions, Motivations, Aspirations, Skills [10–11].
- Intermediate outcomes—sustained problem-related changes over time in participants and others. May include behavior change, normative change, and changes in policies reflect the changes in actions, such as in behaviors and practices, that are a result of increased knowledge and awareness. These are changes in action: Behavior, Procedures, Practice policies, Decisions, Social action [10–11].
- Long-term outcomes—changes in problem-related health measures of the population (e.g. changes in morbidity/mortality) are the conditions that change as a result of actions. Long-term outcomes are what the program is expected to affect. These are changes in conditions as a result of actions: Social, Civic, Health, Environmental [10–11], and are also called Impacts [11].

Table 2.1 provides a crosswalk between components of the CDC Logic Model (first rows) [10, 11] and PIO MM concepts (first column), and shows the congruence of concepts described in intervention effectiveness research, quality improvement activities, and program evaluation. The advantage of the PIO MM over a traditional logic model is the relational structure which specifies how concepts are related. This is essential for developing the analysis plan for the study/evaluation.

2.3 PIO MM and the IHI Quality Improvement Model

The IHI Quality Improvement Model as described in Chapter 1 includes four steps: setting aims, establishing measures, selecting changes, and testing changes [12]. The definitions of these steps are [12]:

- Setting aims. The aim should be time-specific and measurable; it should also define the specific population or system that will be affected [12].
- Establishing measures. Teams use quantitative measures to determine if a specific change actually leads to an improvement [12].
- Selecting changes. Ideas for change may come from those who work in the system or from the experience of others who have successfully improved [12].
- Testing changes. The Plan-Do-Study-Act (PDSA) cycle is shorthand for testing a change in the real work setting—by planning it, trying it, observing the results, and acting on what is learned. This is the scientific method adapted for action-oriented learning [12]. The PIO MM (Fig. 2.1) depicts this phase of the PDSA cycle.

The testing changes step describes the PDSA Cycle that was first used in the 1920s in industry [13]. Careful observation combined with knowledge of scientific evidence underlies the PDSA Cycle, in which a team identifies a problem and decides to do something to test possible changes within the workflow or health system that would improve the problem. Each PDSA cycle consists of rapid hypothesis

Table 2.1 PIO MM Concepts/Logic Model Crosswalk based on the Centers for Disease Control and Prevention Logic Model [10, 11]

	Logic model					
	Process measures			Outcome measures		
PIO MM concepts	Inputs	Activities	Outputs	Short-term	Intermediate	Long-term
Problem (of a population)	Describe problem and the population of interest					
Population context	Tracking of participants and their characteristics	Tracking of response	Population context			
Intervention (activities)	Resources and costs	Actual tasks	Tracking of resources, costs, tasks			
Interventionist	Qualifications, Credentials, training		Fidelity to the planned intervention			
Outcome				Participant changes in problem-related knowledge, awareness, skills, and status	Sustained problem-related changes over time in participants and others	Changes in problem-related health measures of the population
Time	Describe amount			Describe trajectory		

testing based on clinical expertise including knowledge of evidence-based practice. The PDSA cycle requires the team to state the plan (including measures of outcome) and how the plan will be implemented; implement the plan in a small sample; compare measures; seek feedback; and modify the plan for the next cycle. Linking results of sequential PDSA cycles is expected to inform the plan and advance the rapid hypothesis testing to more conditions or settings. When a PDSA cycle leads to demonstration of a desirable outcome, the refined changes may be implemented and spread [13].

The IHI Improvement Model PDSA cycle asserts that the PDSA Cycle is the scientific method adapted for action-oriented learning in the clinical setting [12, 13], and thus it is implied that evidence-based practice or scientific literature underlies the planned change (intervention). The IHI goes beyond evidence-based practice and the literature to suggest that changes may be tried based on good ideas (inspiration) or the experience of others who have successfully improved (practice-based evidence) [12, 13, 14]. This reflects the nature of changes that are needed in real world settings related to diverse workflows and processes for which an evidence-base does not exist. For example, there are many ways to discharge a patient, all of which include numerous important aspects of care that may have grounding in the literature. Insights about the best way to combine elements of discharge planning in a particular practice setting could be tested using the PDSA cycle [13]. However, the impetus for a study examining whether an evidence-based intervention is related to improved patient outcomes or system changes are necessarily based in the scientific literature, and as such, documentation of the rationale for this evaluation should be clearly specified in the PDSA plan. The IHI Quality Improvement model is depicted in a crosswalk with PIO MM in Table 2.2, demonstrating the potential to use the PIO MM in action-oriented quality improvement efforts.

2.4 Using the PIO MM

As we have seen in the PIO MM and crosswalks of the PIO MM concepts with the CDC logic model and the IHI Quality Improvement Model, approaches to demonstrating care quality and outcomes through intervention effectiveness research, quality improvement activities, and program evaluation study must necessarily be comprised at minimum of a problem for which there is an outcome measure and an intervention that can be described or documented. Additional concepts related to setting and interventionist provide context, improving the relevance and interpretability of findings. The PIO MM thus provides a robust meta-model for planning a real-world study, enabling identification of key data and resource needs for completing a successful study [3].

Note that the PIO MM is bounded by the setting within which the population of interest is encountered and intervention(s) occur. All settings have stakeholders who may give feedback regarding the project. Engaging stakeholders creates an

Table 2.2 PIO MM Concepts/IHI Quality Improvement Model Crosswalk based on the Institute for Healthcare Improvement [12]

PIO MM concepts	IHI Quality Improvement Model			
	Setting aims	Establishing measures	Selecting changes	Testing changes
Problem	Addresses an identified health system problem			
Population	Specific population with the problem			
Intervention			Health care quality or system improvement	Trying it (intervention)
Outcome		Quantitative measures to show specific change in the identified problem	Improved patient health (clinical) outcomes that involve both process outcomes (e.g., provide recommended screenings) and health outcomes (e.g., decreased morbidity and mortality).	Observing the results
Time	Time-specific			Real time
Setting			Ideas for change come from those who work in the system	Real work setting

atmosphere of trust, ensures that the study is acceptable, and increases the value of the study findings to the stakeholders. Obtaining feedback from stakeholders can be encouraged by holding periodic discussions, and sharing preliminary findings and draft reports.

Statements that describe and provide rationale for the study based on PIO MM are as follows:

- Statement of the Gap in Knowledge. It is not known if the intervention addressing the problem that is (efficacious or effective) for a given population in (a controlled environment or with other populations in real world settings) also will be associated with positive outcomes in the identified setting and population. This is because there are not studies in the literature specific to the gap we have defined.
- Statement of the Purpose. The purpose of this study is to examine the outcome of the problem after intervention in this setting and population.
- Project Hypothesis. *Interventions* are associated with changes in *Problems* ($P_{Time2} - P_{Time1}$) for the identified *population* in the described *setting* over the given *timeframe*.

- Project Question. Are *interventions* associated with changes in *Problems* ($P_{Time2} - P_{Time1}$) for the identified *population* in the described *setting* over the given *timeframe*?
- Project Goal. Was the *Intervention* associated with changes in *Problems* ($P_{Time2} - P_{Time1}$) for the identified *population* in the described *setting* over the given *timeframe*? Specifying the problem intervention, setting and timeframe, and operationalizing these as measures support the development of SMART goals: Specific, Measurable, Achievable, Relevant and Time-framed [12].

In operationalizing the PIO MM, it is necessary to understand some basic terms related to variables used in PIO MM-based analysis. The following terms used in the results statements are defined here will be further explained and mentioned in Chapter 5 (descriptive statistics), Chapter 6 (inferential statistics), and Chapter 7 (exploratory data analysis):

- Categorical variable—a variable in which valid data include two or more values that are mutually exclusive (categorized) for analysis [15]. Types of categorical variables may include binary variables (two exclusive categories such as high/low or yes/no); nominal variables (scores that have no inherent ordering such as drop-down menu choices for marital status); ordinal variables in which values are inherently ordered such as Likert-type 1–5 scales in which 1 = lowest and 5 = highest).
- Continuous variable—a variable in which valid data include measurements [15]. Types of continuous variables are interval (e.g. temperature), and ratio (any interval measure that starts with 0 meaning there is no value; e.g. age). Temperature in Fahrenheit and Celsius is an interval measures because both scales contain 0. Temperature in Kelvin is a ratio measure that begins after absolute zero [15].
- Dependent variable—a variable that represents the outcome of a statistical model [15].
- Independent variable—a variable that predicts or explains the dependent variable in a statistical model [15].
- Benchmark—a standard value established a priori for measurement of a desired outcome [16]
- Significance—findings that are unlikely to occur due to natural variability in the data (significant) as measured by a *p*-value [15].
- Effect size—indices that describe the size or magnitude of differences between means [17].
- P-value—the probability of obtaining a result at least as extreme as the one observed, if assumptions are met, with a cut off for significance usually established a priori at 5% or 0.05 [15].
- Confidence interval—the interval including the range of likely values for the population within a specified level of confidence (often 95%) [15].
- Mean (average)—the sum of a set of values divided by the number of values [15].

- Standard deviation—a measure of the spread of the data around the mean value for normally distributed data, and is typically reported following a mean [15].
- Standard error of the mean—a measure of the spread of the sample means from repeated samples of a population [15].
- Correlation—the extent to which an association between two variables is linear (may be shown in a straight line) [15].

2.5 Operationalizing the PIO MM

Operationalization of a conceptual framework involves assigning variables to concepts from the model. To operationalize PIO MM, one or more variables must be identified that measure the problem. Ideally, it is useful to also include variables describe the intervention and sample demographics as well as contextual variables such as the attributes of the setting and interventionist:

- Problem measures over time (P_{TimeX}). Quantitative measures of the problem (P) at a point in time (TimeX) must be identified for benchmark or goal attainment, and should be repeated over time to enable comparison of measures (e.g. P_{Time2} to P_{Time1}) and calculate benchmarks.
- Intervention description. Interventions may be described using counts of categorical variables such as terms from classifications or terminologies.
- Interventionist description. Measures of interventionist demographics and qualifications and intervention fidelity may be useful as contextual variables.
- Sample/population characteristics. Demographics of the sample and population including age, race/ethnicity, gender, marital status as applicable.
- Setting characteristics. Variables describing the setting may be continuous as in climate data, or categorical as in census tract data.

The PIO MM, PIO MM-logic model cross-walk, and the PIO MM-PDSA mapping may be used as templates to clarify and define the variables needed to operationalize intervention effectiveness research, quality improvement activities, and program evaluation models [18]. Using the PIO MM enables a systematic check of the concepts that should be operationalized in your project. This process may be helpful in revealing necessary modifications in design and measures during planning in order to clarify the approach and ensure systematic measurement of necessary study concepts [18].

PIO MM measures may be difficult to find, and therefore many studies may omit or provide proxy measures instead of intervention data or other PIO MM concepts. Scholars refer to the Streetlight Effect in which a measure would be selected based on availability (proximity to the streetlight, and therefore visible) rather than the ability to measure the intervention in a meaningful way [19]. All measures selected for PIO MM should be chosen based on their ability to operationalize PIO MM concepts in a meaningful way.

2.6 PIO MM Relationship to Change Theory

The PIO MM depicts what happens when problems are addressed using interventions in real world settings. As a meta-model, the rationale for this change is absent from PIO MM. Such rationale may be explained using an existing change theory. For example, change theories may describe readiness for change or factors related to change [20–24]. Examining change theories may provide rationale for evaluating contextual factors such as participant engagement and stages of change. However, the purpose of change theory is to explain why change happens; thus change theory is explanatory compared to the descriptive nature of the PIO MM; and most change theories fit seamlessly within the PIO MM as supporting literature for the intervention. In other words, a change theory explains why intervention(s) are expected to positively influence problem(s) [20–24]. The theoretical rationale for the intervention is part of the literature review related to the intervention. For example Levine's conservation model describes how the nurse focuses interventions based on the unique characteristics of individuals and their environments in order to help individuals address challenges, adapt as needed, and conserve their personal uniquenesses [24].

In most Intervention effectiveness research, quality improvement activities, and program evaluation studies, it is known that the intervention is efficacious (causes the desired change) and safe. Such research does not aim to examine why the intervention works; rather, change theory is likely known and previously applied in earlier stages of efficacy research. It is possible, but not necessary, to validate change theory as a secondary aim in intervention effectiveness research, quality improvement activities, and program evaluation. To do so, change theory-specific variables should be collected in addition to the PIO MM variables.

2.7 PIO MM Relationship to PICOT

The PICOT Format is a structured approach to defining a research question [25–27]. In the PICOT Format, P stands for Population, I stands for Intervention, C stands for Comparison, O stands for Outcome, and T stands for Time. The PIO MM depicts the concepts of the PICOT Format, and PIO MM may be used in conjunction with PICOT to develop intervention effectiveness research questions [16–18]. Definitions of Population, Intervention, Outcome, and Time are consistent across PIO MM and PICOT. However, the Comparisons concept of PICOT assumes a comparison or control group that receives usual care or no care, and thus expands upon the PIO MM single group before and after design [25–27]. As it may be impossible to include a comparison/control group that receives no intervention, the PIO MM does not require a comparison intervention and differs in this way from the PICOT Format.

While the PICOT question is not reflected in the PIO MM, studies based on the PIO MM may include comparisons of a different sort by segmenting, classifying, and/or stratifying existing datasets and comparing results. Descriptions of the types

of comparisons that can strengthen the single group before and after design follow in Chapter 3, with examples of studies based on PIO MM and comparisons provided in the Part II worksheets.

Reflection Questions

- How does the PIO MM influence your reasoning about your approach to a proposed project?
- Consider the PIO MM concepts in light of your idea: which ones are relevant? Which will vary and which will be constant? Which are likely to be measurable using existing data?
- How can the streetlight effect bias be avoided?

References

1. Gottfredson DC, Cook TD, Gardner FE, Gorman-Smith D, Howe GW, Sandler IN, Zafft KM (2015) Standards of evidence for efficacy, effectiveness, and scale-up research in prevention science: next generation. Prev Sci 16(7):893–926
2. Conn VS, Ruppar TM, Phillips LJ, Chase JA (2012) Using meta-analyses for comparative effectiveness research. Nurs Outlook 60(4):182–190
3. Kernler D. Section 1.2: observational studies versus designed experiments [Internet]. [cited 2017 May 12]. Available from: https://faculty.elgin.edu/dkernler/statistics/ch01/1-2.html
4. Smedley BD, Syme SL (2001) Promoting health: intervention strategies from social and behavioral research. Am J Health Promot 15(3):149–166
5. Paulus JK, Dahabreh IJ, Balk EM, Avendano EE, Lau J, Ip S (2014) Opportunities and challenges in using studies without a control group in comparative effectiveness reviews. Res Synth Methods 5(2):152–161
6. Astleitner H. A systematic approach on the Theoretical Quality of Educational Intervention Research: The Intervention Theory Questions (ITQ). Paper presented at European Conference on Educational Research (ECER), Istanbul, 11–15 September 2013
7. Martin KS (2005) The Omaha System: a key to practice, documentation, and information management, reprinted 2nd ed. Health Connections Press, Omaha, NE
8. dictionary.com. Definition of setting [Internet]. [cited 2017 May 12] Available from: http://www.dictionary.com/browse/setting?s=t
9. dictionary.com. Definition of time [Internet]. [cited 2017 May 12]. Available from: http://www.dictionary.com/browse/time?s=t
10. Silverman B, Mai C, Boulet S, O'Leary L Logic models for planning and evaluation. The Division of Birth Defects and Developmental Disabilities (DBDDD), Centers for Disease Control and Prevention (CDC), Atlanta, GA
11. Krueger RA. Evaluation [Internet]. [cited 2017 May 12]. Available from: http://www.tc.umn.edu/~rkrueger/evaluation.html
12. Institute for Healthcare Improvement. The science of improvement: how to improve [Internet]. [cited 2017 May 12]. Available from: http://www.ihi.org/resources/pages/howtoimprove/scienceofimprovementhowtoimprove.aspx
13. Institute for Healthcare Improvement. IHI Open School. PDSA in everyday life [Internet]. [cited 2017 May 12]. Available from: http://www.ihi.org/Documents/OpenSchoolCourseTranscripts/QI102-Lesson5-Transcript1.html
14. Swisher AK (2010) Yes, exercise is medicine but it is so much more! Cardiopulm Phys Ther J 21(4):4

15. Gay JM. Clinical epidemiology & evidence-based medicine glossary: experimental design and statistics terminology. Washington State University [Internet], 2010. [cited 2017 May 12]. Available from: http://people.vetmed.wsu.edu/jmgay/courses/GlossExpDesign.htm
16. DHHS Agency for Healthcare Quality and Research. Practice facilitation handbook, module 7. Measuring and benchmarking clinical performance [Internet]. [cited 2017 May 12]. Available from: https://www.ahrq.gov/professionals/prevention-chronic-care/improve/system/pfhandbook/mod7.html
17. Kotrlik JW, Williams HA (2003) The incorporation of effect size in information technology, learning, and performance research. Inf Technol Learn Perform 21(1):1–7
18. Centers for Disease Control and Prevention [US] (1999) Framework for program evaluation in public health. MMWR Recomm Rep 48(RR11):1–40
19. Freedman DH (2010) Why scientific studies are so often wrong: the streetlight effect. Discover Mag 26
20. Breuer E, Lee L, De Silva M, Lund C (2016) Using theory of change to design and evaluate public health interventions: a systematic review. Implement Sci 11(1):1
21. Hardeman W, Johnston M, Johnston D, Bonetti D, Wareham N, Kinmonth AL (2002) Application of the theory of planned behaviour in behaviour change interventions: a systematic review. Psychol Health 17(2):123–158
22. Skinner CS, Tiro J, Champion VL (2015) The health belief model. In: Health behavior: theory, research, and practice. Wiley, San Francisco, pp 75–94
23. Glanz K, Bishop DB (2010) The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 31:399–418
24. Schaefer KM, Pond JB (1991) Levine's conservation model: a framework of nursing practice. F.A. Davis, Philadelphia, PA
25. Guyatt G, Drummond R, Meade M, Cook D (2008) The evidence based-medicine working group users' guides to the medical literature, 2nd edn. McGraw Hill, Chicago, IL
26. Riva JJ, Malik KMP, Burnie SJ, Endicott AR, Busse JW (2012) What is your research question? An introduction to the PICOT format for clinicians. J Can Chiropr Assoc 56(3):167–171
27. Stillwell SB, Fineout-Overholt E, Melnyk BM, Williamson KM (2010) Evidence-based practice, step by step: asking the clinical question: a key step in evidence-based practice. Am J Nurs 110(3):58–61

Problem-Intervention-Outcome Meta-Model Project Design

3

3.1 Design for Intervention Effectiveness Research, Quality Improvement Activities, and Program Evaluation

Research design refers to the overall strategy that rationally integrates the components of a project. Adhering to the selected design ensures success in the collection, measurement, and analysis of data according to design and expedites planning [1]. Describing the design will enable all stakeholders to better understand, implement, and/or contribute to your project [2]. There are many ways to describe designs: observational or experimental, prospective or retrospective, comparative or not comparative, cross sectional or longitudinal.

In intervention effectiveness research, quality improvement activities, and program evaluation, a real-world perspective is valued and intended, because complex interactions among individual characteristics, health problems, and interventions may be best understood using data generated in real world programs and healthcare systems. Such projects may be classified as observational studies.

3.1.1 Observational Design

An observational design is intended to describe the characteristics of a population and/or intervention (including outcomes) by studying variables related to individuals in a sample, without manipulation of an intervention or otherwise influencing the variables of interest [3]. Observational design contrasts with experimental design; in which an intervention is manipulated (applied differentially). An experimental design is intended to isolate the effects of the treatment on the outcome variable [3, 4]. The distinction between observational and experimental design may be briefly described as follows: observational studies describe associations or relationships among variables (describe relationships that may be observed), and experimental studies aim to generate new knowledge about cause and effect relationships between variables (learn about cause and effect) [3].

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_3

Observational design necessarily applies to all studies using existing datasets, including studies using electronic health record (EHR) data, census data, and survey data [3]. Therefore observational studies may cover broader populations and settings than experimental studies, and may leverage very large datasets that are becoming available due to technology advances [5]. Projects using big data research methods such as data mining and visualization employ existing datasets and are, by definition, observational studies [6–8]. Such projects seek to identify patterns in the data that may be validated statistically [9–11]. These methods will be further described in Chapter 5 (clustering methods) and Chapter 7 (Exploratory Data Analysis). Furthermore, big data methods known as causal modeling algorithms are emerging to examine causation [12].

Observational design typically applies to quality improvement activities in which an intervention is introduced to a single group (for example in a PDSA cycle) and observations of the process and outcomes are recorded and analyzed [13]. Likewise, observational design applies to program evaluation during which outcome data are recorded as observations of individuals participating in an existing program [2, 14, 15]. In these examples there is a single group that receives a single intervention without manipulation, and therefore no experimental condition exists.

It has been shown that observational studies may yield larger estimates of intervention effectiveness than experimental studies. The reasons for this difference may be several, including the artificiality of the randomized trial with likely reduces the placebo effect for any intervention [16]. Therefore the results of an experimental study may reflect the minimum level of benefit of an intervention, while the observational study may reflect the maximum benefit for groups, while reflecting the integrity of the context in which care was provided [16]. These approaches should be seen as complementary, and each approach should be used as appropriate for the desired purpose.

Intervention effectiveness research, quality improvement activities, and program evaluation therefore are useful and important studies that leverage existing data and/or practice setting activities and programs together with observational design to generate new knowledge, improve practice, and evaluate and guide programs. All of these projects may be retrospective or prospective, and may also be cross sectional (examining an entire sample at one point in time), retrospective (looking back at remembered experience of subjects or examining existing data for particular cases) or prospective (looking forward at the experience of a defined cohort and collecting new data).

3.1.2 Retrospective Design

A retrospective study looks back at existing data or remembered information, and examines independent variables (e.g. individual characteristics and the interventions received) in relation to a dependent variable (outcome of interest). There are numerous threats to the validity of retrospective design due to the lack of ability to

impose controls, and alternative explanations must always be considered. Advantages of retrospective design are the ability to examine rare outcomes in large datasets, and the relatively low cost of using existing data for the analysis [3]:

- Retrospective Cross-sectional Studies describe Associations. A retrospective observational study that uses existing data to examine outcomes at a certain point in time is called a cross-section study. The cross-sectional design helps to understand associations among independent and dependent variables in the entire sample. This may provide essential information upon which to base further study [3].
- Retrospective Case-control Studies enable Comparisons. Case-Control studies are typically retrospective, and may be used to enable comparisons of individuals or groups with and without a particular characteristic. Various methods may be used to establish group membership within existing datasets [3].
- Retrospective Cohort Studies examine Changes over Time. Cohort studies are typically prospective, however, existing data may be used to examine events as they occur over time relative to intervention for groups as in the case control design described above [3].

3.1.3 Prospective Design

A prospective study watches for the occurrence of a future event (e.g. achieving an outcome benchmark). Prospective studies may have fewer threats to validity than retrospective studies, nevertheless it is critical to avoid sources of bias such as the loss of individuals to follow up [3]. Prospective design is suitable if the outcome of interest is relatively common and may be achieve in a short time; however prospective studies may be more costly in terms of resources and time, compared to retrospective design [3].

- Prospective Cross-sectional Studies describe Associations. In prospective studies, the cross-sectional design helps to understand associations among independent and dependent variables in a group of individuals who are available to participate. This may provide essential information regarding recruitment for other prospective studies, including population of interest and power to detect outcomes of interest, upon which to base further study [3].
- Prospective Case-control Studies enable Comparisons. While Case-Control studies are typically retrospective, and the case control design may also be sued to examine and compare individuals or groups with and without a particular characteristic using data collected prospectively after an intervention is implemented, as may occur during quality improvement activities such as PDSA cycles when EHR data are used. As with retrospective case-control studies, various methods may be used to establish group membership within existing datasets [3].
- Prospective Cohort Studies examine Changes over Time. Cohort studies are typically prospective. A cohort study is one in which a group of individuals is

recruited to participate in a study, and the group is observed over a period of time to determine if characteristics of interest may relate to an event or outcome [3]. This process may be lengthy and expensive; however, it may yield true incidence rates and relative risks in comparison with the other designs here described [3].

3.2 Intervention and Measurement Timing

In design terminology, additional descriptions of research designs may be used to specify timing of the interventions and measurements [17]. For example, "before" (or "pre") and "after" (or "post") are terms that refer to measurements taken before and after an intervention is received by a group. The before-and-after design is a common observational design option for intervention effectiveness research, quality improvement activities, and program evaluation projects [17]. The design is intended to demonstrate differences that may be due to intervention for a defined group. Such observational projects may have high external validity because the data often reflect real life situations and groups.

3.3 PIO MM and Research Design

As discussed above, the statistical tests used in a project are determined by the design and the measures that are specified within the design, which is determined by the research question and the available data [3, 4]. The PIO MM depicts the measurement of a problem over time in which an intervention occurs, within the context of a population and setting. At its simplest, it is a *single group before and after design*. This design explicates the basic premise of all intervention effectiveness research, quality improvement activities, and program evaluation projects: that it is possible to effect change in a problem through intervention.

3.4 Benefits and Challenges of the Single Group Before and After Design

There are numerous benefits and challenges of the single group before and after design. In single group before and after design studies, participants are compared with themselves over time, accounting for individual-level variation [17]. However, it is important to note that bias may be introduced due to temporal changes and threats such as regression to the mean, maturation, or multiple contextual or external influences [4, 17]. These circumstances are known as threats to internal validity [4, 17]. This bias in single group before and after studies is unpredictable [4]. Therefore it is essential to compare results with other studies of the same intervention, and to report alternative explanations for the results [4, 17].

3.4.1 Threats to Internal Validity

Threats to internal validity may be understood as alternative explanations for the outcomes or findings of a project. Some common threats to internal validity of observational before and after design include history, regression to the mean, maturation, and drop out; and are described below [4, 17].

- A history threat may occur if one or more events that are unrelated to the intervention take place during the measurement period. Such an event may affect the outcome but would not be reflected in the data [4, 17].
- A regression to the mean threat may occur when the outcome may occur due to the intervention or may be due to chance. The proportion of outcome that may be due to chance would naturally change from extreme values to a more normal value with repeated measurement [4, 17].
- A maturation threat may occur especially over long intervention periods when a group (e.g. infants/children, pregnant women, or older adults) may change due to natural progression of growth/development that are independent of the intervention. Such progression may result in higher or lower outcomes depending on the group, and may be predictable from the data and group characteristics [4, 17].
- A dropout effect threat may occur if some participants are lost or drop out during the measurement period. This may result in differences in the overall characteristics of the group at the time of final measurement. In such a case, outcomes many not be generalizable to the entire sample [4, 17]. It is essential to evaluate the extent of missing data in the sample, as participants with and without a second (final) measure may be missing not at random, and may have been more or less likely to improve, thus biasing the results to either over- or under-estimate intervention effectiveness or outcome [4, 17].

It is critical that these potential threats to internal validity and others that may apply to a particular design/project be reviewed and considered in relationship to the outcomes that are expected and the degree of threat they pose. In doing so, the project may be adjusted during the planning phase, and results may be interpreted appropriately. Supplementing results with complementary information can increase confidence in findings despite one or more threats to internal validity. To validate findings, single group before and after design studies may be compared to statistical or qualitative historical data, or from an external cohort or existing database not drawn from the same institution or population. When working with a large sample, split file cross validation techniques may also be used to examine the internal validity of the findings [18].

3.4.2 Enhancing Before and After Design Using Comparisons

As with case control and cohort designs described above, the before and after design may be enhanced through the use of two or more groups, leveraging measures or

metrics that distinguish groups that exist within the data. These comparison groups may be evaluated as a way of understanding differential intervention effectiveness and outcomes, but unless there is a comparison group that does not receive the intervention, it is not possible to draw causal inferences [17, 19].

3.4.3 Considerations for Prospective Data Collection

For prospective data collection, it may be possible to expand the single group before and after design to include a randomized control group in which the problem is measured for the same population of interest, but no intervention is delivered. This strategy is necessary for efficacy research that aims to show causation of outcome [3, 4, 17]. When it is possible to include a comparison or control group in a prospective intervention study, participants should be randomized, and the study plan should specify exactly how the randomization was done and provide evidence of group equivalence [20]. Due to the resources and time that are needed to conduct randomized trials; this expanded design is beyond the scope of most intervention effectiveness research, quality improvement activities, and program evaluation.

3.5 Comparisons Using PIO MM Variables

The PIO MM concepts operationalized as variables may all be used to create groups within the data in order to enable comparisons: Problem (of a population), Population, Population Context, Intervention, Interventionist, Outcome, and Time.

3.5.1 Problem

The dataset may include one or more problems of interest. If there is more than one problem in the dataset the problem variable enables stratification of all other data to evaluate differential relationships by problem (Fig. 3.1).

3.5.2 Intervention

A dataset may include a number of intervention types that may be analyzed comparatively relative to sample characteristics and outcomes. Intervention data may be binary (yes/no) or may be very granular and datasets resulting from intervention documentation may be very large; requiring preprocessing prior to modeling. Intervention data may include a time measure. Intervention amounts (by counts or total time) may be discretized to identify and group individuals who received varying dosages of the interventions. Interventions may be categorized using definitions or theoretical constructs. Intervention data may be clustered to show co-occurrence or relevance to outcome (Fig. 3.2).

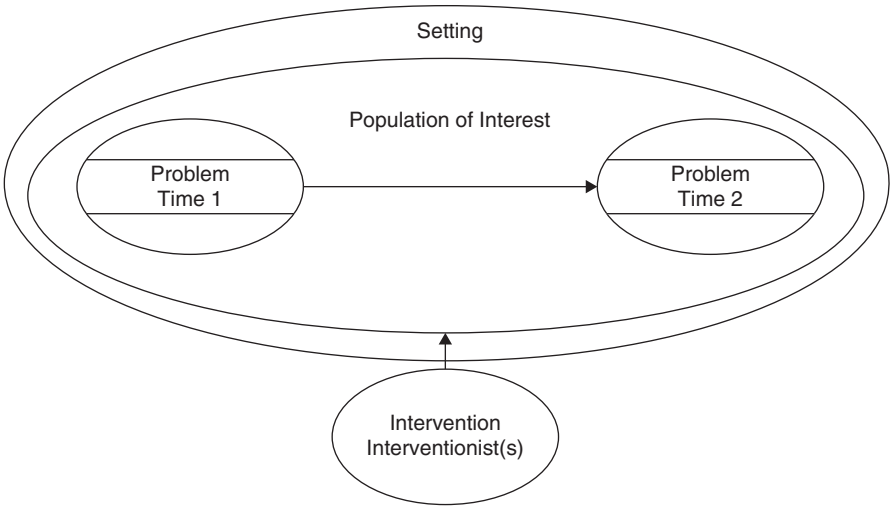

Fig. 3.1 PIO MM with multiple problem variables

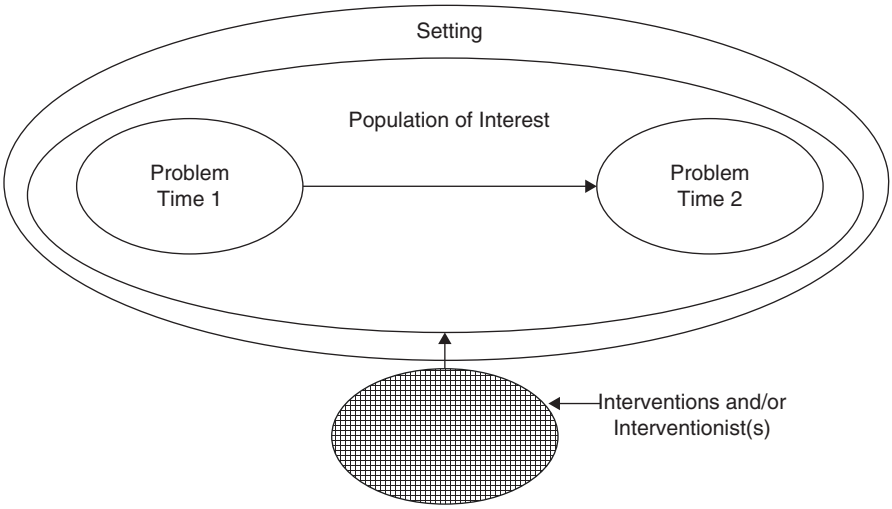

Fig. 3.2 PIO MM with multiple intervention variables and or interventionist variables

3.5.3 Interventionist

Various attributes of the interventionist may be modeled. Variables may include a pseudomized identifier for each interventionist; and/or demographics, credentials, and experience descriptors (Fig. 3.2).

3.5.4 Outcome

A dataset may include a number of outcomes that may be compared (Percentage of sample at P_{Time2} benchmark, Difference of $P_{Time2} - P_{Time1}$, as in Fig. 3.1). Measuring more than one outcome increases interpretability and validity of comparisons. Overall improvement across outcomes should be reported (the individual level), together with specific findings for each problem (the problem-specific perspective). If a before-intervention measurement is not possible, an after-intervention measurement may be compared to a self-described of interventionist-described retrospective before-intervention measurement by asking participants or interventionists to reflect back on their before-intervention assessment or experience of the problem [20].

3.5.5 Population (Individual Characteristics)

Individual characteristics such as demographics and baseline assessments may be used to stratify or classify a sample into groups for comparative analysis (as with retrospective case control or retrospective cohort designs).

3.5.6 Setting

Setting characteristics such as geographic jurisdiction or health system, clinic, unit, or variables that describe these characteristics may be used to stratify individuals or interventions (Fig. 3.3).

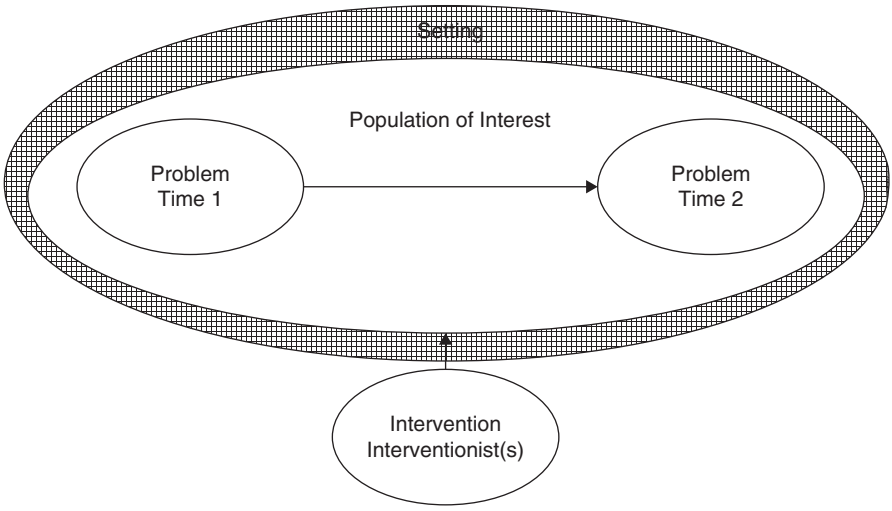

Fig. 3.3 PIO MM with multiple setting variables

When groups are created within a large dataset using stratification or other methods, statistical analysis of group effects must be conducted for each group to which intervention effects are associated. Analysis of between-group differences should be conducted and reported, and significant group differences should be accounted for in outcome models [19]. Further in-depth exploration may look at associations between interventions and outcomes, and models of outcome interactions. When multiple outcomes are analyzed, the researcher must provide a clear rationale for the treatment of multiple outcomes which may be [19].

3.5.7 Time

The PIO MM assumes use of a before and after design and measurement strategy. The study plan should specify timing of all measurements. For studies collecting data prospectively, it is prudent to establish times or intervals in time for the before-intervention and after-intervention measurements. A length of stay or duration of treatment variable may be computed by computing the differences between the before and after measurement times [20]. Duration of treatment may have an important influence on outcomes. Short and long-term effects may be analyzed if there are at least three or more measurements over time; ideally with the final measurement taking place well after the intervention is over [19]. Multiple measurements of problems or interventions over time may be visualized to show changes over time and incorporated in longitudinal models. If time associate with each intervention is available, these increments and the total time of interaction may be modeled similarly.

3.6 Mixed Methods: Qualitative Evaluation

It is helpful to triangulate quantitative outcome findings using qualitative analysis of interviews or other narratives related to the intervention. In such interviews or surveys, use open-ended questions to ask what is different or what has changed, and what caused the change. The questions should relate to variables of interest in the quantitative analysis. A self-assessment question using a Likert-type scale (e.g. 1 = low and 5 = high), asking participants to rate the problem now versus before the intervention may also provide useful data for comparison. It is also instructive to ask participants to share a story that captures their experiences [21]. Such data may be analyzed using thematic analysis to provide new insights that may support or extend the findings of quantitative analysis.

Operationalizing the design depends on simple (and sometimes complex) tools and resources that are described in Chapter 4. Design will guide selection and use of these resources.

Reflection Questions

- What is the value and/or importance of design and how will design reflect and/or guide your project?
- How can an observational design be strengthened to increase confidence in the findings of projects using existing data?
- Which threats associated with observational designs apply to the project? How can they be addressed?

References

1. Trochim WM. The research methods knowledge base, 2nd edn. [Internet]. [cited 2017 May 12]. Available from: http://www.socialresearchmethods.net/kb/ (version current as of October 20, 2006).
2. Centers for Disease Control and Prevention [US] (1999) Framework for program evaluation in public health. MMWR Recomm Rep 48(RR11):1–40
3. Kernler D. Section 1.2: observational studies versus designed experiments [Internet]. [cited 2017 May 12]. Available from: https://faculty.elgin.edu/dkernler/statistics/ch01/1-2.html
4. Silverman B, Mai C, Boulet S, O'Leary L (2009) Logic models for planning and evaluation. The Division of Birth Defects and Developmental Disabilities (DBDDD), Centers for Disease Control and Prevention (CDC), Atlanta GA
5. Hersh WR, Cimino J, Payne PR, Embi P, Logan J, Weiner M, Bernstam EV, Lehmann H, Hripcsak G, Hartzog T, Saltz J (2013) Recommendations for the use of operational electronic health record data in comparative effectiveness research. EGEMS 1(1)
6. Tukey JW (1977) Exploratory data analysis. Addison-Wesley, Reading
7. Monsen KA, Peterson JJ, Mathiason MA, Kim E, Lee S, Chi CL, Pieczkiewicz DS (2015) Data visualization techniques to showcase nursing care quality. Comput Inf Nurs 33(10):417–426
8. Monsen KA, Brandt JK, Brueshoff B, Chi CL, Mathiason MA, Swenson SM, Thorson DR (2017) Social determinants and health disparities associated with outcomes of women of childbearing age who receive public health nurse home visiting services. J Obstet Gynecol Neonatal Nurs 46(2):292–303
9. Mayer-Schönberger V, Cukier K (2013) Big data – a revolution that will transform how we live, work and think. John Murray Publishers, London
10. Hey T, Tansley S, Tolle K (2009) The fourth paradigm: data intensive scientific discovery. Microsoft Research, Redmond, WA
11. Schneeweiss S (2014) Learning from big health care data. N Engl J Med 370(23):2161–2163
12. Aliferis CF, Statnikov A, Tsamardinos I, Mani S, Koutsoukos XD (2010) Local causal and markov blanket induction for causal discovery and feature selection for classification part i: algorithms and empirical evaluation. J Mach Learn Res 11(Jan):171–234
13. Institute for Healthcare Improvement. The science of improvement: how to improve [Internet]. [cited 2017 May 12]. Available from: http://www.ihi.org/resources/pages/howtoimprove/scienceofimprovementhowtoimprove.aspx
14. Rural Health Information Hub. Types of evaluation in health promotion and disease prevention programs [Internet]. [cited 2017 May 12]. Available from: https://www.ruralhealthinfo.org/community-health/health-promotion/4/types-of-evaluation
15. World Health Organization (1998) Health promotion evaluation -recommendations to policy-makers. Report of the WHO European working group on health promotion evaluation. WHO, Copenhagen
16. Black N (1996) Why we need observational studies to evaluate the effectiveness of health care. Br Med J 312(7040):1215

17. Before-and-after design: A simple evaluation design [Internet]. [cited 2017 May 12]. Available from: http://158.132.155.107/posh97/private/research/evaluation/Chp_3.pdf
18. Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD (2001) Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 54(8):774–781
19. Paulus JK, Dahabreh IJ, Balk EM, Avendano EE, Lau J, Ip S (2014) Opportunities and challenges in using studies without a control group in comparative effectiveness reviews. Res Synth Methods 5(2):152–161
20. Krueger RA. Evaluation [Internet]. [cited 2017 May 12] Available from: http://www.tc.umn.edu/~rkrueger/evaluation.html
21. Erickson KJ, Monsen KA, Attleson IS, Radosevich DM, Oftedahl G, Neely C, Thorson DR (2015) Translation of obesity practice guidelines: measurement and evaluation. Public Health Nurs 32(3):222–231

Tools for Intervention Effectiveness Research, Quality Improvement Activities, and Program Evaluation

4

This chapter provides an overview of the necessary resources for successful completion of intervention effectiveness research, quality improvement activities, or program evaluation in real-world settings based on the PIO MM.

4.1 Data Sources

Sources of data for Intervention effectiveness research, quality improvement activities, and program evaluation vary widely, but all of the variables from diverse data sources can be mapped to the PIO MM as described in Chapters 2 and 3. This will enable analyses of outcome ($P_{Time2} - P_{Time1}$; % attaining a pre-defined P_{Time2} benchmark) in relationship to the identified sample's demographic characteristics and interventions received over time within the context of the situation and location.

This book emphasizes the benefits of using existing data sources to complete your project. The availability of data from EHRs and other sources is growing exponentially and there are extensive but largely untapped possibilities for reusing existing data [1–5]. When it is necessary to collect new data, keep in mind that additional resources must be available to support the data collection process.

The advent of computer technology in health care and in everyday life, and the resulting availability of large amounts of existing data, is causing a shift in thinking about health care research in which the data and data mining or machine learning techniques are used to generate hypotheses for researchers to investigate [20–22]. Hey and colleagues called this Big Data shift the *fourth paradigm of research* [2], building on thousands of years of empirical research (first paradigm), centuries of theoretical research (second paradigm), and decades of simulation research (third paradigm) [2]. While many big data methods are highly

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_4

complex, this book provides basic information about simple Exploratory Data Analysis (EDA) techniques that may help detect patterns within data to inform intervention effectiveness research, quality improvement activities, and program evaluation. EDA may be used in Big Data/fourth paradigm projects to depict data graphically in order to detect outliers, trends and patterns; and to suggest hypotheses for further investigation [6].

Mayer-Schönberger and Cukier describe three major shifts in data use for research that are occurring due to the use health information technology [1]. The first is the ability to analyze vast amounts of data about a topic rather than be forced to settle for smaller sets. This gives us the ability to examine health-related patterns of large groups as well as drilling into patterns for small groups identified within the large datasets. The second is a willingness to embrace data's real-world messiness rather than privilege exactitude as we are taught to do in the empirical and theoretical sciences. This means that there are conflicting notions regarding the value of Big Data in scientific inquiry, and calls for further development of criteria for rigorous Big Data research. The third is a growing respect for correlations rather than a continuing quest for elusive causality. This supports the notion that understanding and describing what is happening may be as important as understanding why something happens. In the complex real-world of health care, this is an important perspective given that it is nearly impossible to control for all of the factors that influence health outcomes [1].

From a statistical perspective, having more data means less sampling error. However, the data are necessarily more challenging to manage because of the size and variability across systems. Data management resources and skills are essential for success, and often a team approach is preferred. Furthermore, lack of controls around data entry by the many health care clinicians and assistants who are entering data into EHRs. Thus, it is necessary to examine the data for improbable values, make decisions about missing data, and otherwise account for the biases inherent within such observational data [1–3, 7, 8].

4.2 Checklists for Obtaining New or Existing Data for Operationalizing the PIO MM

Two checklists are provided here to enable comparison of the resources and skills needed to obtain existing and new data (Tables 4.1 and 4.2).

Table 4.1 Obtaining existing data

• Identify study design and sample including case controls within the sample.
• Identify one or more sources of existing data that that measure the problem using continuous variable(s) (ordinal, interval, or ratio)
• Determine sample size (power analysis) needed to achieve a medium effect in the analysis of $P_{Time2} - P_{Time1}$ for the instruments
• Identify one or more sources of existing data that describe interventions
• Identify a means of documenting intervention fidelity using the existing data or agency training protocols
• Identify any variables that describe the sample, interventionist, and context within the existing data, consider any comparison groups that may be identified using these variables.
• Ensure that existing data are de-identified or, if identified, that permissions are in placed to use identified data (e.g. permission from agency director)
• Confirm exemption from review by institutional review board for reuse of existing data without identifiers
• Store data without identifiers (use fake IDs for all individuals and interventionists in the sample) in secure location.

Table 4.2 Obtaining new data

• Identify study design and sample. If there will be a comparison group, incorporate the rationale and consenting process for the comparison as well as intervention groups in all following steps.
• Identify one or more instrument(s) to measure the problem using continuous variable (ordinal, interval or ratio) considering operationalization of the problem from more than one perspective, and minimizing respondent burden
• Determine sample size (power analysis) needed to achieve a medium effect in the analysis of $P_{Time2} - P_{Time1}$ for the instruments
• Identify instrument(s) to capture intervention description
• Identify a means of documenting intervention fidelity
• Identify characteristics of the sample, interventionist, and context that will support the analysis
• Identify participants who are best able to provide the data: Interventionists, individuals receiving care, other observers, or a combination
• Interviews of participants may be used to obtain qualitative data to confirm, add nuance, and/or negate quantitative findings.
• Obtain permission from institutional review board to collect the data from the population of interest
• Enroll participants using standard consent procedures and collect data at desired intervals.
• Store data without identifiers (use fake IDs for clients and interventionists) in secure location.

4.3 Electronic Health Record Data

The expected and desired goal of electronic health record (EHR) implementation beyond data capture for care management is the reuse of data for evaluation and research in real time and for population health improvement [7–8]. Electronic health record data is a potential data source for Intervention effectiveness research, quality improvement activities, and program evaluation [18–19]. Ideal EHR data are accessible, standardized, and taxonomic, including multiple problem measurements over time (P_{Time1}, P_{Time2} … P_{TimeX}); intervention documentation over time; and client characteristics (e.g. demographics). Large EHR datasets are advancing the notion of data exploration as a new paradigm of research that leverages the power of the data to uncover patterns that may be related to intervention effectiveness and outcomes [1–3, 7, 8].

Accessing EHR data or other databases within the health system may require the permission of management or leadership, and the skills of the EHR/informatics team. A university or consortium may have EHR data or other data sources that may be accessed through secure data shelters. The processes by which data may be accessed will vary by institution and dataset. It is imperative to begin the process early in order to facilitate successful completion of the project in a timely manner. The data that are available will likely define one or more aspects of the design and methods used in your project; therefore it is of primary importance to identify and understand the data sources that will be most beneficial for your project.

4.4 Nursing-Specific Data

Nursing scholars have long been leaders in the development of interface terminologies for healthcare. Beginning in the 1970s, researchers have advanced the science of describing, classifying, and using defined terms for assessment, diagnosis, intervention, and outcome measurement for the purpose of knowledge representation, interoperability, and reuse of clinical data. These efforts have resulted in robust tools that are described below [9–20].

The Nursing Management Minimum Data Set consists of core essential data needed to support the administrative and management information needs for the provision of nursing care. The standardized format allows for comparable nursing data collection within and across organizations for nurse and health system characteristics, and nurse and health system credentials [13].

Nursing Minimum Data Set: The NMDS is a system of describing nursing interventions using a minimum set of elements of information with uniform definitions and categories concerning the specific dimensions of nursing, which meets the information needs of multiple data users in the health care system. Components of the NMDS are standardized terminologies for documenting demographic characteristics and outcomes; and nursing assessments and interventions [10–12]. This enables comparison to other intervention descriptions, revealing the contents of the Black Box [17] (the condition of having received complex interventions) and

explicating tailored protocols. Such terminologies are intended for care planning, documentation, communication of care across settings, integration of data across systems and settings, effectiveness research, and other health care information management functions [9–16]. The four intervention terminologies (CCC [16], ICNP [14], NIC [15] and the Omaha System [9]) described below can be used within the PIO MM for highly granular intervention descriptions.

Of these, the Omaha System is uniquely suited for PIO MM because it is a comprehensive ontology for health care that enables relational documentation of PIO MM variables including problem characteristics (signs/symptoms), and problem measures over time (P_{TimeX}), and problem-specific interventions [9]. Furthermore, it exists in the public domain and is therefore readily available to researchers, clinicians, and evaluators. This book provides exemplars using the Omaha System that may be generalized to any other classification or terminology, including reference terminologies such as SNOMED CT [21] and LOINC [22], in which the interface terminologies are embedded [9, 12, 14–16].

Clinical Care Classification (CCC) System. The CCC is a Standardized Coded Nursing Terminology for the electronic documentation of Clinical Nursing Practice. It consists of a four–level framework/hierarchy: 1st level) 4 Healthcare Patterns, 2nd level) 21 Care Components (Classes), 3rd level) Nursing Terminologies: 176 Nursing Diagnoses, 804 Nursing Interventions/Actions, and 528 Nursing Outcomes, and 4th level) 3 Outcome Qualifiers & 4 Action Type Qualifiers [16].

International Classification of Nursing Practice (ICNP). The ICNP is a standardized nursing terminology included within acknowledged terminologies of the World Health Organization. The ICNP has seven axes describing different areas of nursing and related interventions, enriched by two special axes related to pre-coordinated Diagnosis/Outcomes (DC) and Operations (IC) which facilitate daily use in practice [14].

Nursing Interventions Classification (NIC). The NIC, first published in 1997, is a comprehensive standardized classification of interventions nurses perform. The NIC defines an intervention as any treatment based upon clinical judgment and knowledge that a nurse performs to enhance nurse-sensitive outcomes. There are 554 Interventions and nearly 13,000 activities in the 2013 edition [15]. The NIC is often used in combination with the diagnoses described by the North American Nursing Diagnosis Association-International (NANDA-i) [18] and the outcomes described by the Nursing Outcomes Classification (NOC) [19–20].

4.5 Omaha System

Omaha System consists of three relational, reliable, and valid components designed to be used together: the Problem Classification Scheme (problem assessment; signs/symptoms), Intervention Scheme (interventions describing care plans and services), and the Problem Rating Scale for Outcomes (problem change/evaluation). Initial developmental, reliability, validity, and usability research was conducted during 4 federally-funded projects between 1975 and 1993. A second edition was published in 2005 [9].

4.5.1 Problem Classification Scheme

The Problem Classification Scheme provides a structure, terms, and system of cues and clues for a standardized assessment of individuals, families, and communities. It helps practitioners collect, sort, document, classify, analyze, retrieve, and communicate health-related needs and strengths. It is a comprehensive, orderly, non-exhaustive, mutually exclusive taxonomy or hierarchy in three levels: Domains, Problems, and signs/symptoms. The first level describes a comprehensive, ecological view of health in four Domains: Environmental, Psychosocial, Physiological, and Health-related Behaviors. Forty-two health problem concepts are at the second level; each Problem concept is taxonomically classified within one of the Domains. By definition, Problem concepts are intended to be neutral to enable specification of strengths and challenges related to each concept. Clusters of signs and symptoms (s/sx) that describe actual problems are unique to each Problem concept and provide binary (yes/no) items that may be aggregated for analysis (e.g. sum of all s/sx or selected s/sx) [9].

4.5.2 Intervention Scheme

The Intervention Scheme is a comprehensive, orderly, non-exhaustive, mutually exclusive taxonomy in a three level hierarchy that is used in conjunction with the Problem Classification Scheme. The first or most general level consists of four actions (called Categories) that may be used to address any of the 42 Problems: (1) Teaching, Guidance, and Counseling, (2) Treatments and Procedures, (3) Case Management and (4) Surveillance. The second level describes 75 objects of action (called Targets) that further specify the intervention. At the care description level, guideline, protocol, or care description information may be used to guide and document evidence-based care specific to a population, practice, or policy [34]. Thus each intervention consists of three defined terms (Problem, Category, and Target), and a customizable care description. This hierarchical structure enables analysis across levels of granularity (time-specific intervention counts of P, C, T; and any combinations thereof) and evaluation of intervention fidelity (delivered relative to expected intervention frequencies) [9].

4.5.3 Problem Rating Scale for Outcomes

P_{Time1}, P_{Time2}…P_{TimeX} The Problem Rating Scale for Outcomes is a method to evaluate progress throughout the period of service. It consists of three five-point, Likert-type scales to measure the entire range of severity for the concepts of Knowledge (K_{Time1}, K_{Time2}…K_{TimeX}), Behavior (B_{Time1}, B_{Time2}…B_{TimeX}), and Status (S_{Time1}, S_{Time2}…S_{TimeX}). Knowledge is d efined as what the person knows; Behavior as what the person does; and Status as the number and severity of the person's signs and symptoms or predicament. Each of the subscales is a continuum providing an evaluation framework for examining Problem-specific ratings at regular or predictable times. Suggested times for measurement include admission, specific interim points, and dismissal [9]. Figures 4.1 and 4.2 provide examples of Problem Rating Scale for Outcomes data in spreadsheet and Comma Separated Values (CSV) formats, respectively.

Fake ID	Problem	KTime1	BTime1	STime1	KTime2	BTime2	STime2
6908D	Nutrition	2	4	3	4	5	5
6908D	Sleep and rest patterns	3	4	3	4	4	5
6908D	Abuse	2	4	3	5	4	5
6908D	Medication regimen	3	4	3	4	5	5
55081S	Medication regimen	3	4	3	3	4	5
55081S	Pregnancy	2	4	3	3	4	5
55081S	Sanitation	3	4	3	4	4	5
55081S	Residence	2	5	3	3	5	5
55081S	Bowel function	3	3	4	4	5	5
55081S	Income	2	3	4	4	2	3
55081S	Nutrition	3	3	4	4	2	3
2236P	Nutrition	3	3	3	4	4	4
2236P	Personal care	3	1	3	5	2	4
2236P	Health care supervision	2	3	4	4	2	4
2236P	Neighborhood/workplace safety	2	2	4	4	2	4
2236P	Circulation	2	2	3	3	4	4
2236P	Personal care	2	2	3	3	2	2
2236P	Pain	2	2	4	4	1	4
2236P	Income	2	2	3	3	1	4
2236P	Skin	2	2	4	4	2	4

Fig. 4.1 Example of a spreadsheet with fake ID, Problem, and Problem Rating Scale for Outcomes Variables at two time points (Time1 and Time2) for Knowledge (K), Behavior (B), and Status (S)

```
FakeID,Problem,KTime1,BTime1,STime1,KTime2,BTime2,STime2
6908D,Nutrition,2,4,3,4,5,5
6908D,Sleep and rest patterns,3,4,3,4,4,5
6908D,Abuse,2,4,3,5,4,5
6908D,Medication regimen,3,4,3,4,5,5
5081S,Medication regimen,3,4,3,3,4,5
55081S,Pregnancy,2,4,3,3,4,5
55081S,Sanitation,3,4,3,4,4,5
55081S,Residence,2,5,3,3,5,5
55081S,Bowel function,3,3,4,4,5,5
55081S,Income,2,3,4,4,2,3
55081S,Nutrition,3,3,4,4,2,3
2236P,Nutrition,3,3,3,4,4,4
2236P,Personal care,3,1,3,5,2,4
2236P,Health care supervision,2,3,4,4,2,4
2236P,Neighborhood/workplace safety,2,2,4,4,2,4
2236P,Circulation,2,2,3,3,4,4
2236P,Personal care,2,2,3,3,2,2
2236P,Pain,2,2,4,4,1,4
2236P,Income,2,2,3,3,1,4
2236P,Skin,2,2,4,4,2,4
```

Fig. 4.2 Example of a the same data in comma separated values (CSV) format with fake ID, Problem, and Problem Rating Scale for Outcomes Variables at two time points (Time1 and Time2) for Knowledge (K), Behavior (B), and Status (S)

4.6 Analysis Software and Techniques

Analysis of PIO MM variable data for intervention effectiveness research, quality improvement activities, and program evaluation may be accomplished using EDA and standard descriptive and inferential statistics that can be computed using basic spreadsheet functions. Specialized statistical software may be used depending on your project needs and the skills of the investigator/evaluator and team members.

4.7 Power Analysis

Decisions about the amount of data that are needed to have sufficient power for statistical calculations are made based on power analysis. Power analysis should be conducted in advance, and should describe the sample size needed to detect effects of an intervention based on clinical considerations regarding the expected outcome [21–23]. There are numerous nuances that depend on the project purpose. Therefore, it is optimal for clinicians to work with a statistician to estimate the range of sample size that is appropriate for a given project. For existing data, the estimated sample size is known, and power calculations may be used to verify that the proposed analysis using the sample will have sufficient power to detect effects. For new data, power calculations will offer guidance for recruitment of a sample size that is neither too big (and thus wastes resources) or too small (and thus will not detect effects of the intervention) [21–23].

4.8 Software for Descriptive and Inferential Statistical Methods and for Creating Graphs/Charts

Many options are available for managing, analyzing, and depicting PIO MM data. Four common statistical software resources are described as examples of the functionality that is needed to analyze data and generate tables and charts to display findings: Excel, R, SAS, and SPSS [24–27].

4.8.1 Microsoft Excel

Microsoft Excel is a commonly available spreadsheet software (part of Microsoft Office) that includes statistical functions [24]. There is no additional cost for Excel within the Microsoft Office suite, and numerous statistics functions are embedded within the software with user-friendly prompts that enable a broad array of statistical analyses. Directions for using Excel are freely available in Microsoft Office Help functions and also on-line through various internet chat rooms and videos. Basic or advanced Excel skills are useful for data management and are related or transferrable to skills needed for other statistical packages and programs.

4.8.2 R

R is open source statistical software that is freely available. The R Project for Statistical Computing is a software environment for statistical computing and graphics [28]. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. R is an environment within which statistical techniques are implemented. R provides a wide variety of statistical (e.g. linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering) and graphical techniques, and is highly extensible. Users need to learn code (the R language) to work with the program. Directions for using R may be offered through academic statistics courses and are also available through free on-line tutorials. Copy-ready R coding examples for many diverse analyses are freely available on-line.

4.8.3 SAS

SAS (Statistical Analysis System) is a software suite developed by SAS Institute for advanced analytics, multivariate analyses, business intelligence, data management, and predictive analytics. There is a free version for students and teachers [26]. Users need to learn programming code specific to SAS in order to work with the program. Directions and tutorials for using SAS may be offered through academic statistics courses and are also freely available through the company and on-line. SAS is commonly used by epidemiologists and biostatisticians as well as statistical experts in many fields.

SPSS. SPSS (Statistical Package for the Social Sciences) is a software owned by IBM [25, 27]. The point and click interface makes it easy to use for multivariate analyses, business intelligence, data management, and predictive analytics. The SPSS Software has a basic 2-screen layout. IBM® SPSS® Student GradPack is a single-user license that provides affordable access to the SPSS comprehensive software package with essential tools for statistical analysis, modeling, and data mining research. Directions and tutorials for using SPSS may be offered through academic statistics courses and are also freely available through the company and on-line. SPSS does not require learning code, but the coding syntax behind the point and click interface is available and may be saved for re-use.

4.9 Big Data (Pattern Detection) Methods

Techniques and methods for detection of patterns are important in intervention effectiveness research, quality improvement activities, and program evaluation [1–8, 29]. The above described statistical software as well as specialized software below may be used for detection of patterns using methods such as clustering and visualization analysis. Discussion of specialized environments such as noSQL and Hadoop to manage and analyze big data are beyond the scope of this book [29]. A few of the tools that are available for clustering and data visualization techniques are described below. The use of these techniques is discussed further in Chapters 5 and 7.

4.9.1 Clustering

Weka Data Mining Software in Java. Weka is a collection of machine learning algorithms for data mining tasks. The algorithms can either be applied directly to a dataset or called from your own programming using Java code. Weka contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also well-suited for developing new machine learning schemes. Weka is open source software issued under the GNU General Public License. There are several free online courses that teach machine learning and data mining using Weka [30–31].

MATLAB. MATLAB® is used to analyze and design systems and products and is incorporated within automobile active safety systems, interplanetary spacecraft, health monitoring devices, smart power grids, and LTE cellular networks. It is used for machine learning, signal processing, image processing, computer vision, communications, computational finance, control design, and robotics. MATLAB can be accessed from mobile devices or any web browser. MATLAB uses matrix based natural computational mathematics expressions for ease in coding and maintenance. The MATLAB language also provides features of traditional programming languages, including flow control, error handling, object-oriented programming, unit testing, and source control integration. MATLAB student-use software includes the same functionality and is available to students through the MATLAB web site [32].

4.9.2 Visualization

D3. D3.js or Data Driven Documents is a JavaScript library for manipulating documents based on data; combining powerful visualization components and a data-driven approach [33]. D3 allows users to bind arbitrary data to a Document Object Model (DOM), and then apply data-driven transformations to the document. For example, D3 may be used to generate an HTML table from an array of numbers, or, with the same data it may be used to create an interactive SVG bar chart that incorporates smooth transitions and interaction. D3 enables efficient manipulation of documents based on data. It is open source and uses java script. D3 is extremely fast, supporting large datasets and dynamic behaviors for interaction and animation. D3's functional style allows code reuse through a diverse collection of components and plugins [33].

Tableau. Tableau is a live visual analytics tool for data exploration, with interactive dashboards to aid in discovery of patterns. Tableau enables rapid image building and calculations from existing data using drag-and-drop processes that can be combined with statistical methods such as trend analyses, regressions, and correlations. Interactive maps may be created automatically using postal codes recognized from over 50 countries. An Academic Programs version of Tableau for students may be downloaded without charge from the Tableau Web site [34].

4.10 Team Approach

Health care is a complex social intervention that increasingly generates data for reuse in intervention effectiveness research, quality improvement activities, and outcome analysis [4–5]. The PIO MM provides a simple conceptualization of healthcare interactions for operationalization by variables that enable modeling of inter-relationships among variables and overall outcomes following intervention. Success in modeling is optimally achieved through combining skill sets from disciplines including the health professions (content expertise related to the population, context, problem, intervention, and expected outcomes), statisticians (power analysis, design, analytical skills, and guidance regarding interpretation of findings), as well as method-specific expertise for studies employing big data methods (e.g. clustering, visualization). Research, quality improvement, and evaluation teams leverage the skills of many disciplines to achieve the results using these skills with the tools described here, and others.

Putting the team together with the tools and resources (including data) leads to knowledge discovery using a wide array of calculations and other methods of synthesizing information. Chapter 5 begins with methods for preparing data and leads into descriptive and data mining methods.

Reflection Questions

- What are the advantages and disadvantages of using new data? Existing data?
- What software do you use that will enable you to complete the project? What skills will you need to learn?
- Considering the data and software requirements of your project, describe your ideal team and the roles of each team member.

References

1. Mayer-Schönberger V, Cukier K (2013) Big data – a revolution that will transform how we live, work and think. John Murray Publishers, London
2. Hey T, Tansley S, Tolle K (2009) The fourth paradigm: data intensive scientific discovery. Microsoft Research, Redmond, WA
3. Schneeweiss S (2014) Learning from big health care data. N Engl J Med 370(23):2161–2163
4. Raghupathi W, Raghupathi V (2014) Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2(1):3
5. Kruse CS, Goswamy R, Raval Y, Marawi S (2016) Challenges and opportunities of big data in health care: a systematic review. JMIR Med Inform 4(4)
6. Tukey JW (1977) Exploratory data analysis. Addison-Wesley, Reading, MA
7. Olsen L, McGinnis JM (2010) Redesigning the clinical effectiveness research paradigm: innovation and practice-based approaches: workshop summary. National Academies Press, Washington, DC

8. Jensen PB, Jensen LJ, Brunak S (2012) Mining electronic health records: towards better research applications and clinical care. Nat Rev Genet 13(6):395–405
9. Martin KS (2005) The Omaha System: a key to practice, documentation, and information management, reprinted 2nd edn. Health Connections Press, Omaha, NE
10. Werley HH (1991) Nursing minimum data: abstract tool for standardized comparable, essential data. Am J Public Health 81(4):421–426. doi:10.2105/AJPH.81.4.421
11. Sewell JP, Thede LQ. Nursing and informatics: opportunities and challenges. Nursing documentation in the age of the EHR [Internet], 2012. [cited 2017 May 12]. Available from: http://dlthede.net/informatics/Chap16Documentation/chap16.html
12. American Nurses Association [ANA]. ANA recognized terminologies that support nursing practice [Internet], 2012. [cited 2017 May 12]. Available from: http://www.nursingworld.org/MainMenuCategories/Tools/Recognized-Nursing-Practice-Terminologies.pdf
13. Huber D, Delaney C (1997) The American Organization of Nurse Executives (AONE) research column. The nursing management minimum data set. Appl Nurs Res 10:164–165
14. Marucci AR, De Caro W, Petrucci C, Lancia L, Sansoni J (2015) ICNP-International Classification of Nursing Practice: origin, structure and development. Prof Inferm 68(2):131
15. Butcher HK, Bulechek GM, Dochterman JM, Wagner C (2013) Nursing interventions classification (NIC). Elsevier Health Sciences
16. Clinical care classification [Internet]. [cited 2017 May 12]. Available from: https://www.sabacare.com/about/
17. Nielsen K, Randall R (2013) Opening the black box: presenting a model for evaluating organizational-level interventions. Eur J Work Organ Psy 22(5):601–617
18. Herdman TH (ed) (2011) Nursing diagnoses 2012-14: definitions and classification. John Wiley & Sons
19. Moorhead S (2013) Nursing outcomes classification (NOC). Elsevier Health Sciences
20. Johnson M (ed) (2001) Nursing diagnoses, outcomes, and interventions: NANDA, NOC, and NIC linkages. Mosby Incorporated
21. Cohen J (1992) A power primer. Psychol Bull 112(1):155–159
22. Centers for Disease Control and Prevention [US] (1999) Framework for program evaluation in public health. MMWR Recomm Rep 48(RR11):1–40
23. Röhrig B, du Prel JB, Wachtlin D, Kwiecien R, Blettner M (2010) Sample size calculation in clinical trials: part 13 of a series on evaluation of scientific publications. Dtsch Arztebl Int 107(31-32):552
24. Mendenhall W, Beaver RJ, Beaver BM (2012) Introduction to probability and statistics. Cengage Learning
25. Green SB, Salkind NJ (2010) Using SPSS for windows and Macintosh: analyzing and understanding data. Prentice Hall Press
26. SAS University edition [Internet]. [cited 2017 May 12] Available from: http://www.sas.com/en_us/software/university-edition.html
27. IBM® SPSS® Student GradPack [Internet]. [cited 2017 May 12]. Available from: http://www-03.ibm.com/software/products/en/spss-stats-gradpack
28. What is R? [Internet]. [cited 2017 May 12]. Available from: https://www.r-project.org/about.html
29. Tiwari V, Tiwari B, Thakur RS, Gupta S (2013) Pattern and data analysis in healthcare settings. Google Books
30. Weka Project http://www.cs.waikato.ac.nz/ml/weka/index.html
31. Frank E, Hall MA, Witten IH (2016) The WEKA workbench. In: Online Appendix for "Data mining: practical machine learning tools and techniques", 4th edn. Morgan Kaufmann
32. MATLAB [Internet]. [cited 2017 May 12]. Available from: https://www.mathworks.com/products/matlab/features.html#matlab-speaks-math
33. D3 data driven documents [Internet]. [cited 2017 May 12]. Available from: https://d3js.org/
34. What is Tableau? [Internet]. [cited 2017 May 12]. Available from: http://www.tableau.com/products/desktop#TGETWeLmfvacBMZt.99

Descriptive Analysis and Interpretation

5

5.1 Introduction

Data cleaning and pre-processing are approaches that help to understand and prepare for analysis, and must be completed after obtaining data, and before proceeding to analysis. There is a saying among data analysists that for any data-driven project, data analysts spend 80% of their time cleaning and pre-processing the data, and the other 20% of the time complaining about cleaning and pre-processing it. This is to emphasize the importance of ensuring that the data are organized according to the data analysis plan, with no known improbable or erroneous values, and are ready to be employed in your project. After the data cleaning and pre-processing phase is complete, the actual modeling of the data may take very little time. The guidance provided here is for existing datasets. For new data entry, additional processes apply, including double data entry and data checking processes [1].

5.2 Data Cleaning

Data cleaning is an essential first step in dealing with existing data, or an important next step after data entry for projects using new data [1]. This is a systematic project that includes actively searching for implausible values by sorting and examining the data using various descriptive methods for identifying outliers (as well as inliers that are erroneous data points falling within range) and values that are not within the realm of real life possibility. Examples may include a score outside of the range of a measure (e.g. less than 1, or greater than 5, on a scale of 1–5); values generated by different units of measure (e.g. temperature of 98.6 Celsius would not be compatible with life, but in Fahrenheit it is normal human body temperature). Data should be compared based on other characteristics in the dataset. For example, a height of 26 inches (66 cm) would be a reasonable height for an infant or toddler (ages 1–2 years) but would be much less likely for an adult (aged 18 or greater). Similarly, data entry errors may be caused by entering data in reverse

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_5

fields (such as entering length and head circumference of a newborn). Such an error may be detected systematically for a subset of newborns that appear to be very short for gestational age but on further investigation are found to have disproportionately large head circumferences. The data cleaning process is iterative and likely will require several iterations to discover and resolve issues. Even after data are cleaned, other errors may be detected during or after analysis by the investigator/evaluator based on experience, expected results, previous studies, evidence in the literature, or common sense [1]. Data cleaning may be approached in three phases: the Screening Phase, Diagnostic Phase, and Treatment Phase [1].

5.2.1 Screening Phase

Detecting errors and inconsistencies in data should be approached systematically through the use of simple descriptive analysis described below using standard statistical packages or spreadsheets as described in Chapter 4. Useful screening methods may include:

- Creating data tables describing means, ranges, and standard deviations, and checking them for unusual or unexpected values
- Creating frequency tables to identify of variables and/or values that are out of range or inconsistent with expected values
- Depicting distributions using graphs such as histograms, box plots, and/or scatter plots
- Comparing variable combinations using cross-tabulations to identify inconsistent or implausible associations [1]

5.2.2 Diagnostic Phase

To clarify whether unusual data points, patterns, and statistical results of the sample identified in the Screening Phase can be classified as being one of five types:

- Erroneous—a false or impossible value
- True extreme—at the extreme edge of expected value as established a priori
- True normal—the expectation established a priori was incorrect
- Idiopathic—not obviously false, but no explanation found or connection to other data available

Both hard (strict) and soft (flexible based on criteria) cut offs and clinical judgment are needed. It is helpful to examine other related data for coherence across values (for example, low weight scores may be confirmed with a diagnosis of malnutrition). Descriptive and inferential statistics and visualization techniques may be helpful in clarifying these patterns [1].

5.2.3 Treatment Phase

In this phase, decisions are made about correcting, deleting, or leaving unchanged any errors that were confirmed in the Diagnosis Phase. Heuristics for the Treatment Phase include the following:

- Impossible values should be deleted or corrected if possible based on other known values.
- True extreme values should be kept in the data.
- If truly erroneous data constitute a significant part of the sample it may be necessary to start over with a new dataset.
- Any implausible data that are eliminated should be treated as missing data [1–3].

5.2.4 Missing Data

Frequency analysis may be conducted to document missing data [2]. It is important to understand the extent of missingness, and whether or not data are missing at random. If possible, missing data should be identified, calculated, or inferred using a variable that is equivalent and available in another source [3]. Data missing not at random may bias findings, and thus missing data of substantive proportions must be considered when interpreting findings [2–3]. It is important to know how missing data is coded within the dataset. There may be a placeholder for missing data, such as the number 9999, or missing data may be a blank cell that may be indistinguishable a cell left empty because a value that did not apply and therefore was not documented. Missing data may inform the study in a number of ways and should be examined for meaningful patterns [2–3].

5.3 Pre-Processing

5.3.1 Transforming and Recoding

In order to successfully model studies based on PIO MM, it is often desirable to create new variables from existing variables. For example, pre-processing may needed in order to specify group membership for case control design, discretize a continuous P_{Time2} measure to benchmark outcomes, or create a new metric from combinations of variables that may together provide a more accurate or complete indicator of the PIO MM concepts or population groups. Several examples of pre-processing techniques are provided below, including ways of transforming and recoding various PIO MM concepts. New, transformed, recoded, or computed variables must be evaluated for errors, improbable values, and missing data as described above [1]:

- A new variable may be created to transform a continuous variable to a categorical variable. For example, a variable called "Adult" may be created based on a continuous "Age" variable by assigning a value of 1 to every individual aged 18 years or older, and a 0 to every person aged less than 18 years. This provides a single variable that enables stratification of the sample by adults (0 = no, 1 = yes) and also enables evaluation of the model separately for individuals who are or are not yet adults.
- This same technique may be used to create a variable that denotes whether a minimum benchmark (value for P_{Time2}) was attained (0 = no/1 = yes).
- A new variable may be created by calculating a change score by subtracting $P_{Time2} - P_{Time1}$. This is easily achieved in most spreadsheet programs using simple formula functions. The difference between these two existing variables becomes a P_{Change} score that may then be analyzed as an outcome variable along with P_{Time2}.

Multiple variables may be combined into a single metric such as a risk score or other metric to discriminate between risk groups using numerous variables [4–6]. Below are examples of metrics developed using PIO MM variables including a Social and Behavioral Determinants Index [4], a Maternal Risk Index [5], and a Hospitalization Risk Score [6]:

- Social and Behavioral Determinants of Health (SBDH) Index [4]. A SBDH metric was developed based on the IOM-recommended SBDH measures using existing s/sx and demographic data. The IOM-recommended SBDH domains mapped to 19 items. These were summed to create the SBDH Index. The SBDH Index metric was used to stratify a sample into groups with increasing SBDH item counts for comparative program evaluation.
- Maternal Risk Index (MRI) [5]. A MRI variable was computed for each individual using weighted totals of high impact problems, adjusted by baseline knowledge. The MRI scores were partitioned at the median of the distribution to form low- and high- risk individuals. The MRI was used in modeling individual risk relative to intervention tailoring and intervention effectiveness [5, 7].
- Hospitalization Risk Score (HRS) [6]. The HRS was developed theoretically based on the literature before viewing the sample data to identify factors that predicted hospitalization for individuals receiving care at home for acute or chronic health conditions. Candidate predictors summarized from the literature review were finalized using clinical expert review, mapped the final predictors to problem terms, and weighted the problems based on literature review. The HRS was used to examine the ability to predict re-hospitalization in EHR data.

5.3.2 Identification and Labeling of Clusters Within a Sample

Techniques such as clustering may be used to pre-process large datasets in order to determine group membership for individuals in the sample, or to classify interventions into meaningful clusters. Clusters may then be used as variables to model groups of individuals or groups of interventions in further hypothesis testing or

program evaluation [8–16]. Several examples of methods used for clustering are provided from the literature:

- Intervention data may be clustered using probabilistic algorithms [8–9]. Unsupervised clustering (clustering without conditioning on an outcome variable) may be used in order to independently examine the relationships between interventions and outcomes in further analysis [10]. Using this approach, novel clusters of co-occurring home care interventions were discovered using multiple iterations of K-Means and expectation maximization [8–10]. Cluster meanings were interpreted by clinical experts based on frequencies of the interventions in the machine-generated clusters. The intervention clusters were used in a model of intervention effectiveness for prevention of hospitalization outcomes [11].
- Intervention data may also be clustered using Multilevel Partitioning algorithms [12]. Novel clusters of co-occurring interventions were discovered using KMETIS Multilevel Partitioning algorithms and a weighted undirected intervention graph to an assign of the interventions to unique clusters [12–13]. Preprocessing was needed to assign interventions to a single co-occurring intervention and create the intervention graph [13]. Cluster meanings were again interpreted by clinical experts based on cluster content. The intervention clusters were used to examine intervention effectiveness for improving individual outcomes [13].
- Individuals with similar signs/symptoms may also be clustered using a variational expectation maximization algorithm [14–16]. The final result was a mixed membership assignment of individuals to groups, and then classifying individuals by selecting the group to which an individual has the highest membership. Hence, the final output contains clustering of similar individuals who have similar characteristics as well as the final estimated distribution over characteristics for each group. Because initialization of the model is randomized, it can generate different final clusterings on different runs. True patterns in the data will persist over multiple runs [13]. Cluster meanings were again interpreted based on cluster content [13].

5.4 Descriptive Statistics

A descriptive statistic is a numerical value that summarizes the values in a sample [17]. As discussed earlier in this chapter, descriptive techniques are essential for data cleaning in preparation for a project. When implementing intervention effectiveness research, quality improvement activities, and program evaluation, descriptive statistics are used to summarize and describe the characteristics of variables in a given sample or dataset to provide a clear understanding of central tendencies, distributions, and proportions within the sample. Particularly in relationship to the PIO MM, it is critical to be able to describe all aspects of the PIO MM concepts (Chapter 2) in order to create a credible platform for use of inferential statistics relating PIO MM concepts (Chapter 6) and use of EDA to detect patterns in the data (Chapter 7). Definitions of terms are provided below for frequencies and ranking, measures of central tendency, and measures of distribution.

5.4.1 Frequency

The number of times a value occurs in a dataset is its frequency. The frequency analysis is most often used with categorical variables. A frequency is often reported together with a percentage to help interpret the meaning of the number, as in 66 (27%) of individuals in the sample smoked cigarettes; or 122 (51%) were overweight or obese. For binary variables, report the number and percent of the most frequent result as in 233 (80%) of individuals in the sample were female; or 27 (56%) attained the desired benchmark [17]. Bar charts are useful depictions of the relative frequencies of categorical variables. Examples of PIO MM variables that may be analyzed using frequency and percent include:

- Problem, if more than one: all problems of interest in the project
- Intervention, if more than one: interventions used to address the problems during the study
- Interventionist: credentials and/or categorical demographics of the interventionist
- Outcome: binary and categorical outcomes if applicable
- Population: categorical demographic characteristics of the sample
- Setting: categorical unit, health system, geographic location, or jurisdiction characteristics

Interpretation of frequencies is limited to which values occur more or less often in a dataset. Figure 5.1, shows which problems occur with more or less frequency in the sample by arranged counts from largest to smallest.

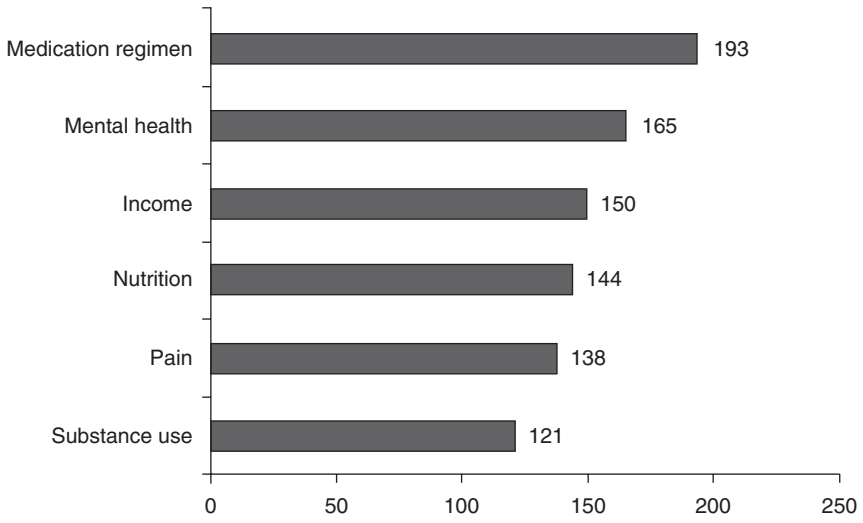

Fig. 5.1 Bar chart showing frequencies of problems

5.4.2 Cross Tabulation (Cross Tab) Matrix

Cross tabulation is a method of displaying frequency data in order to show comparisons between two or more variables [18]. The relative frequencies of values for one variable compared to another variable may be summarized in a two (or more) dimensional table in which number (frequency) of each cell describes shared populations for both variables by variable value. Interpretation of cross tabs involves a Chi-Square (χ^2) test of the significance of expected cell frequencies compared to actual cell frequencies, which will be described in Chapter 6 [18].

5.4.3 Rank

The relative position of a numeric value compared to other values in the dataset is its rank. Ranks may be reported by percentiles (e.g. 75th percentile), or the data value that separates the values above and below the rank value that closest to the corresponding percentage of the sample. The median (50th percentile) divides the sample at the midpoint value. Dividing a sample by rank and frequency results in groups of similar size between cut offs called quartiles (for 4 groups that each comprise roughly a fourth of individuals in the sample), quintiles (5 groups), deciles (10 groups), and so forth. These groups may vary in size based on the distribution of individuals in the sample. Likewise, the cut off points may be unevenly distributed in the range of the variable. A box and whisker plot may be used to portray distribution by rank in percentiles (Fig. 5.2) [17].

Interpretation of ranked data enables comparison that is normalized across variables using percentiles. Figure 5.2 shows the largest change in status overall for the Skin problem. The median value of change in status was also highest for Skin, followed by Mental health, and then Pain.

5.4.4 Measures of Central Tendency

The median, mean, and mode are three measures of central tendency. The mean, or average, is the sum of the values divided by the number of values. The median, described above, may be higher or lower depending on the distribution of all other values in the dataset. The mode is the most frequent value and may differ from the mean and median, and when graphed, shows how the distribution may vary with one or more peaks (Fig. 5.3) [17].

Interpretation of measures of central tendency for the distribution of the variable for children and adolescents ages 1–21 years clearly differs for the mean (age 9), median (age 11), and mode (age 5) (Fig. 5.3). This means that the most children were 5 years old, but there is a bi-modal distribution with the second most children at age 12. These known measures may guide division of the sample into two groups (ages 1–8 and ages 9–21) to reflect the distribution; absent knowledge of any program factors or relevant clinical details that would suggest other ways of splitting the sample.

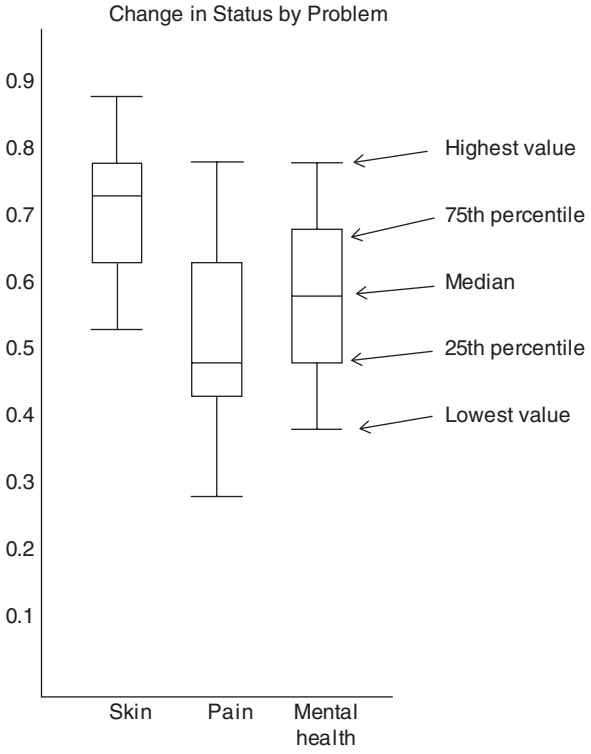

Fig. 5.2 Box plot representing 25th, 50th, and 75th percentiles for values of three variables. *Whisker lines* show the highest and lowest values

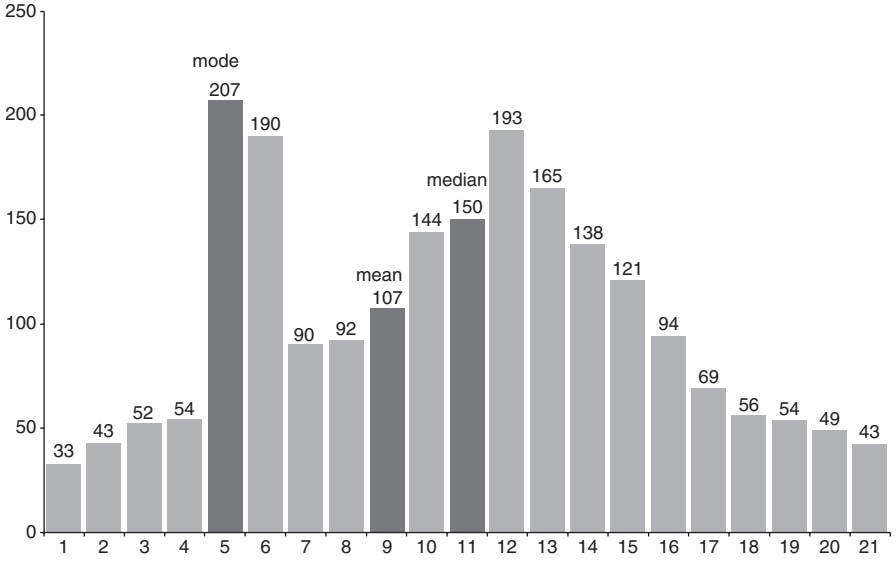

Fig. 5.3 Measures of central tendency of values in a distribution

5.4.5 Measures of Distribution

In addition to central tendency measures, a distribution of values as shown in Fig. 5.3 may also be described in several ways: a simple range (smallest value to largest value) and other calculated measures: outliers, standard deviation (SD), and standard error of the mean:

- Outliers are a value larger or smaller than the median plus or minus 1.5 times the interquartile range. These may be indicated in a box plot by dots or asterisks.
- Standard deviation is a measure of the spread of the data around the mean value for normally distributed data, and is typically reported following a mean.
- Standard error of the mean is a measure of the spread of the sample means from repeated samples of a population. The SEM may be reported after a mean.

Results statements based on PIO MM that incorporate descriptive statistics are summarized in Table 5.1 as follows:

Table 5.1 Results statements for descriptive analyses, with standard reporting conventions

• **Description of sample characteristics:** *The characteristics of the sample were* (XX%) (categorical characteristics), (*M* = XX.XX, *SD* = X.XX) (continuous characteristics).
• **Description of interventions:** *There were XXX interventions, with an average of* (*M* = XX.XX, *SD* = X.XX) *per* (person, visit, problem). *Most interventions were for the* (problem, group) (XX%).
• **Description of outcomes:** *On average, final outcomes* (P_{Time2}) *for the entire sample were* (*M* = XX.XX, *SD* = X.XX), *an average change of* ($P_{Time2} - P_{Time1}$) (*M* = XX.XX, *SD* = X.XX) *from the baseline score* (P_{Time1}) (*M* = XX.XX, *SD* = X.XX)
• **Description of benchmark attainment:** *Of the entire sample*, XX% *attained the desired benchmark of* XX (state the definition of the benchmark number).

The descriptive statistical techniques discussed here will provide the foundation for understanding the scope and substance of intervention effectiveness research, quality improvement activities, and program evaluation projects; as well as building blocks for further inferential analysis (Chapter 6) to compare variables, and exploratory data analysis (EDA) (Chapter 7) to reveal the patterns in the data.

Reflection Questions

- Which descriptive statistics would be important in the data-cleaning phase of your project? Why?
- What data computations or transformations would be useful to help summarize concepts as a new variable in your project? What are the formulae or algorithms for creating the variable?
- In using descriptive statistics for the PIO MM variables, which would be most critical in understanding the population and its problems, interventions, and outcomes?

References

1. Van den Broeck J, Cunningham SA, Eeckels R, Herbst K (2005) Data cleaning: detecting, diagnosing, and editing data abnormalities. PLoS Med 2(10):e267
2. Smith AK, Ayanian JZ, Covinsky KE, Landon BE, McCarthy EP, Wee CC, Steinman MA (2011) Conducting high-value secondary dataset analysis: an introductory guide and resources. J Gen Intern Med 26(8):920–929
3. Pigott TD (2001) A review of methods for missing data. Educ Res Eval 7(4):353–383
4. Monsen KA, Brandt JK, Brueshoff B, Chi CL, Mathiason MA, Swenson SM, Thorson DR (2017) Social determinants and health disparities associated with outcomes of women of childbearing age who receive public health nurse home visiting services. J Obstet Gynecol Neonatal Nurs 46(2):292–303
5. Monsen KA, Radosevich DM, Kerr MJ, Fulkerson JA (2011) Public health nurses tailor interventions for families at risk. Public Health Nurs 28(2):119–128
6. Monsen KA, Swanberg HL, Oancea SC, Westra BL (2012) Exploring the value of clinical data standards to predict hospitalization of home care patients. Appl Clin Inform 3(4):419–436
7. Monsen KA, Peterson JJ, Mathiason MA, Kim E, Lee S, Chi CL, Pieczkiewicz DS (2015) Data visualization techniques to showcase nursing care quality. Comput Inform Nurs 33(10):417–426
8. Witten IH, Frank E (2005) Data mining: practical machine learning tools and techniques. Morgan Kaufmann
9. Tan PN, Steinbach M, Kumar V (2005) Introduction to data mining. Addison-Wesley, Reading
10. Monsen KA, Westra BL, Yu F, Ramadoss VK, Kerr MJ (2009) Data management for intervention effectiveness research: comparing deductive and inductive approaches. Res Nurs Health 32(6):647–656
11. Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. J Mach Learn Res 3(Jan):993–1022
12. Karypis G, Kumar V (1999) Parallel multilevel series k-way partitioning scheme for irregular graphs. SIAM Rev 41(2):278–300
13. Monsen KA, Westra BL, Oancea SC, Yu F, Kerr MJ (2011) Linking home care interventions and hospitalization outcomes for frail and non-frail elderly patients. Res Nurs Health 34(2):160–168
14. Monsen KA, Banerjee A, Das P (2010) Discovering client and intervention patterns in home visiting data. West J Nurs Res 32(8):1031–1054
15. Shan H, Banerjee A (2008) Mixed-membership naive Bayes models (technical report TR 09–002). Department of Computer Science & Engineering, University of Minnesota, Twin Cities
16. Wang H, Shan H, Banerjee A (2009) Bayesian cluster ensembles. In: SIAM international conference on data mining. Society for Industrial and Applied Mathematics, Philadelphia, PA. pp 209–220.
17. Gay JM. Clinical epidemiology & evidence-based medicine glossary: experimental design and statistics terminology. Washington State University [Internet], 2010. [cited 2017 May 12]. Available from: http://people.vetmed.wsu.edu/jmgay/courses/GlossExpDesign.htm
18. Dunn OJ, Clark VA (2009) Basic statistics: a primer for the biomedical sciences. Wiley

Inferential Analysis and Interpretation

6

6.1 About Inferential Statistics

Inferential statistics are used to test whether or not differences are likely due to chance for patterns discovered in descriptive analysis (Chapter 5) and EDA (Chapter 7). Statistical significance means that the findings, such as differences in the mean benchmark attainment between groups, would be unlikely to occur due to natural variability in the data [1]. The study design, variables, and sample sizes that are available in the dataset determine the questions that may be addressed, hypotheses that may be tested, and the statistical techniques that are appropriate [2–3]. There are numerous approaches that may be appropriate, and disciplinary perspectives from biostatistics, epidemiology, and educational psychology (and many others) may influence selection of the tests that are optimal in each proposed project. It is wise to consult with mentors and statistical support services in planning the models that will best estimate the differences that are important in intervention effectiveness research, quality improvement activities, and program evaluation. This handbook offers basic guidance for selected approaches that will enable rigorous analysis of comparisons for PIO MM variables; it is not intended to be a comprehensive statistical textbook. For a more details it is advisable to access statistical resources available in universities, other institutions and agencies, and on-line.

After cleaning and pre-processing, data may be evaluated to determine which tests are appropriate and may be used to evaluated differences and associations. In general, statistical approaches should be selected based on the types and distributions of values for the PIO MM variables, and the expected fit of the variables with the parametric or non-parametric assumptions [3]. Recall that variables may be categorical (binary, nominal, or ordinal) or continuous (see definitions in Chapter 2). Further terms that are used to describe distributions and statistical tests are presented here: Normal distribution, Parametric test, Non-parametric test:

- Normal distribution: a bell-shaped curve shown in a histogram that displays all data plotted as frequencies by value for any given variable. This shape arises

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_6

from random variation. Normal distribution is one of the assumptions for the use of parametric statistical tests [1–3].
- Parametric Test: a statistical test that yields a statistic and p-value, that is based on mathematical measures of distribution variability such as standard deviation or standard error (see definitions in Chapter 2) [1].
- Non-parametric Test: a statistical test that uses the relative positions or ranks (as defined in Chapter 5) of the data and yields a statistic and a p-value [1].

Deciding whether to use parametric and non-parametric tests. The assumptions that underlie and support the use of parametric statistics and non-parametric statistics are important when deciding which test to use [2–3]. Some of the most frequent assumptions for use of most parametric statistical tests are noted below. This is not a comprehensive list, but rather a general idea of assumptions that indicate that a data sample is well suited for use with parametric tests:

- Variable values are interval or ratio scale of measurement
- random sampling from a defined population
- normal distribution
- equal variances if two or more groups or variables in the design

Regarding the first assumption, the requirement for interval or ratio data for use of parametric tests, it is important to note that the scores of many valid, reliable instruments are ordinal scales. Such scores are typically treated as continuous variables when meeting other assumptions. Some statisticians disagree on the use of parametric statistics with such ordinal scales [2–3] and advocate for purity in the use of non-parametric statistics; however, these do not allow for comparison of the central tendencies which may limit inferences regarding correlations among variables (mean, S.D.) [2–3].

Regarding the second assumption, random sampling from a defined population, retrospective analysis of existing large datasets from EHRs or registries may be analyzed with the understanding that the aim is to capture data from an entire population. Typically, inferential tests are used in order to infer or estimate relationships among attributes of a population using a representative sample; but in the case of large datasets, the actual outcomes of the entire population may be present, and inference would not be necessary. Random sampling techniques may be incorporated within retrospective models to simulate randomization. For example, this may be achieved by randomly splitting a sample for case control design (Chapter 3), or repeated analysis (e.g. clustering algorithms and visualization techniques) to understand patterns that persist across the entire sample (Chapter 5). This does not eliminate the selection bias that occurs with data from a health system or agency in which individuals who did not receive services cannot be included in the analysis.

Regarding the third assumption, normal distribution will not be perfect but should be unimodal and fairly symmetrical as determined by visually examining histograms of the data (see Fig. 6.1). There are tests that evaluate goodness of fit

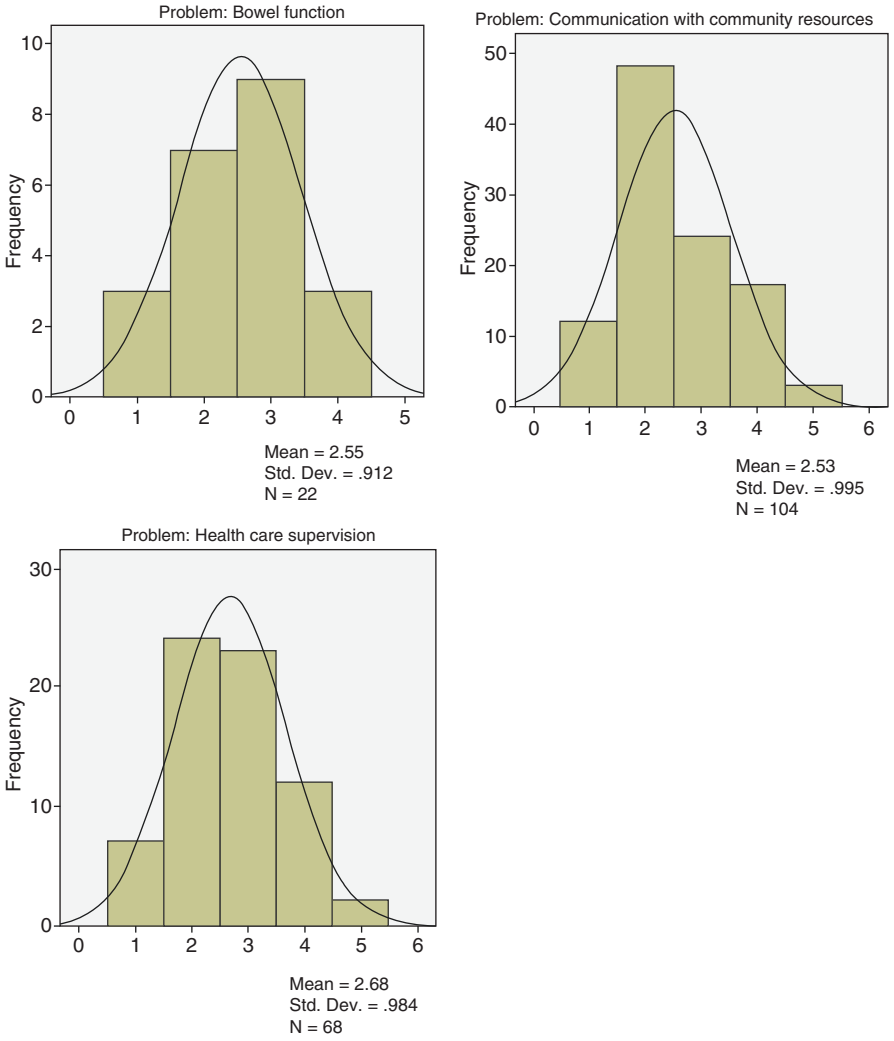

Fig. 6.1 Histograms showing normal curve overlay on data distribution for Likert-type ordinal scales rating health problems (1 = lowest, 5 = highest)

with particular distributions, such as the Anderson-Darling and Shapiro-Wilk tests [4]. Distributions that are not normal may be transformed using mathematical algorithms, however this should be considered with caution, as the results of tests using a transformed distribution may be difficult to interpret [2–3].

Regarding the fourth assumption, variances may be evaluated during the data cleaning and descriptive statistics phases (Chapter 5) to enable decision making regarding use of parametric or non-parametric tests.

6.2 Comparisons and Statistical Significance

6.2.1 Comparisons of Sample Characteristics

In Chapter 5 the descriptive statistical tests commonly used to report characteristics of the sample were discussed. When there is analysis by group in a project, it is helpful to provide a comparison of sample characteristics between groups. Thus, the means and standard deviations for the continuous data would be reported by group, together with a statistical test of differences between the means, such as a t-test which is commonly used to compare means [2]. To fully report t-test results the following format is recommended: t(XX) = X.XX, p =. XX [5] in which (XX) is the degrees of freedom. Similarly, the frequencies and percentages of categorical variables would be reported by group, together with comparisons of differences in frequencies often using cross tabs with a χ^2 (Chi-Square) statistic [2]. To fully report Chi-Square results the following format is recommended: χ^2 (X, N = XX) = 0.XX, p =. XX, in which (X, N = XX) are the degrees of freedom and sample size [5].

Interpretation of Sample Characteristics State overall sample characteristics, followed by any significant between group differences, with statistic and *p-value*. For example, *there were 155 adults in the sample, ages 17–44, with an average age of 24.8 (5.4) years. The majority were female (60%) and white (68%) or African American (30%); with 20% of the total sample identifying as Hispanic ethnicity. Compared to females, males were significantly older* $t(90) = 5.43$, $p < 0.001$, *and more likely to be white* $\chi^2(1, N = 94) = 0.89$, $p = 0.035$. Note that this type of reporting assumes that only significant differences between groups need to be reported.

6.2.2 Outcomes as Measured by Before and After Comparison

The PIO MM difference in outcome computed using change in outcome scores over time is a paired samples problem ($P_{change} = P_{Time2} - P_{Time1}$). Paired samples tests such are used to evaluate significance of the change for a *single group and outcome*. Common paired samples tests are the t-test (parametric) in which means are compared; or Wilcoxon signed-rank test (non-parametric) in which ranks are compared [2–3]. To fully report Wilcoxon signed-rank test the following format is recommended: Z = X.XX, p =. XX [5].

Interpretation of Before and After Comparisons State whether change in the problem after intervention did or did not occur, and was or was not significant. For example, *knowledge of nutrition improved significantly following PHN intervention* ($t(366) = 4.42$, $p < 0.001$). If there is a control group it is possible to infer causation due to the intervention. If there is no control group, it is possible to know that interventions may be associated with the outcome, but not to infer causation. In addition, it is important to state alternative explanations that may have contributed to the findings. For example, *it is possible that knowledge increased due to multiple factors in addition to the interventions received during PHN visits, including but not limited to*

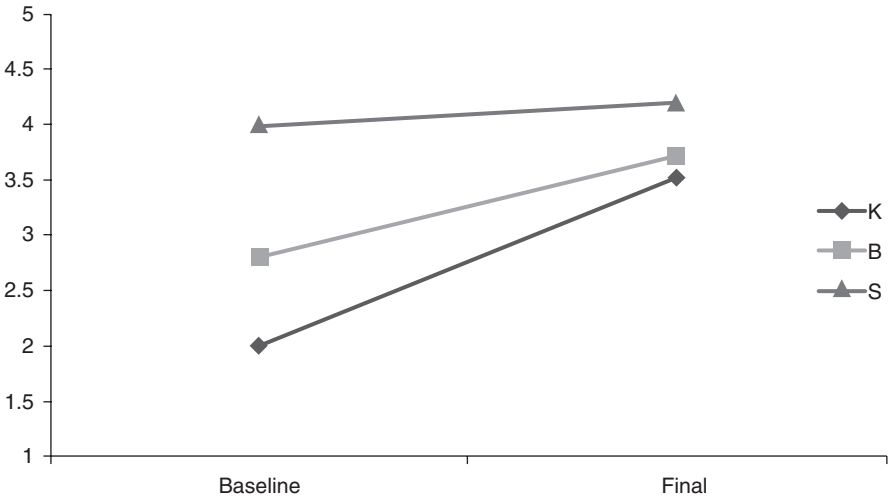

Fig. 6.2 Line graphs depicting change in Knowledge, Behavior, and Status scores for the Nutrition problem by showing averaged baseline and final scores. Note that Knowledge scores change the most (steepest slope) but have the lowest final scores, while Status scores change the least but have the highest final scores. Note that comparisons are facilitated using the mean (averaged) scores

nutritionist information after PHN referral to the Women, Infants, and Children (WIC) supplemental nutrition program, or maturation due to the length of PHN services (Fig. 6.2).

Variations among and between groups relative to before and after comparisons may be detected using analysis of variance (ANOVA) with post hoc analysis (parametric) or a Kruskal–Wallis test (non-parametric) [2–3, 6]. To fully report ANOVA main effects and interaction results the following format is recommended: F(X, XXX) = X.XX, p =. XX, and a significant interaction, F(X, XXX) = X.XX, p =. XX) [5]. To fully report Kruskal-Wallis H results the following format is recommended: H(X) = X.XX, p = .XX where (X) is the Chi-Square value [5].

Interpretation of Between Group Comparisons State whether there were significant between group differences in outcome after intervention. *Analysis of variance showed a main effect of intervention group on outcomes (*$F(1, 144) = 5.41$, $p = 0.02$*1) and post hoc analysis using Bonferroni that outcomes were significantly lower for group A compared to all other groups (*$p = 0.03$).

In addition, it is important to state alternative explanations that may have contributed to the findings. For example, between group differences may be influenced by factors other than the demographic variables used to stratify groups such as the season of the year in which the different groups received the intervention.

Variations among and between groups relative to before and after comparisons may be further examined by statistically controlling for the effects of other continuous variables such as individual age and number of signs/symptoms using ANCOVA (parametric) or Mantel-Haenszel Chi-square (nonparametric) [2–3, 6]. To fully report ANCOVA results, use the recommended format described above for ANOVA.

6.2.3 Benchmarking

Benchmarking [7, 8] (e.g. $P_{Time\ 2}$ compared across groups) is an evaluation approach used to analyze and compare outcomes that was used first at the Xerox Corporation and defined as "finding and implementing best practices" [8, p. 230]. Benchmarking enables a comparison of interventions or outcomes to an established standard. This promotes the identification and sharing of best practices and catalyzes change and quality improvement. Mixed methods models have been used to compare population characteristics and benchmark across counties [8]. Within and between least squares means ratings were compared, and *differences of differences* in ratings were reported across the groups using corresponding *p-value*s [8]. Benchmarking is especially useful in evaluating the level of outcome attainment, and should be reported with change scores ($P_{change} = P_{Time2} - P_{Time1}$) to provide a more complete picture than reporting change alone or final outcomes alone.

Interpretation of Benchmarking State whether there were significant between group differences in outcome after intervention. For example, *overall, 80% of the population of interest reached the benchmark of 4 (adequate knowledge) for the Nutrition problem. Individuals in group E were more likely than the other counties to attain this benchmark* ($p = 0.012$) (Fig. 6.3).

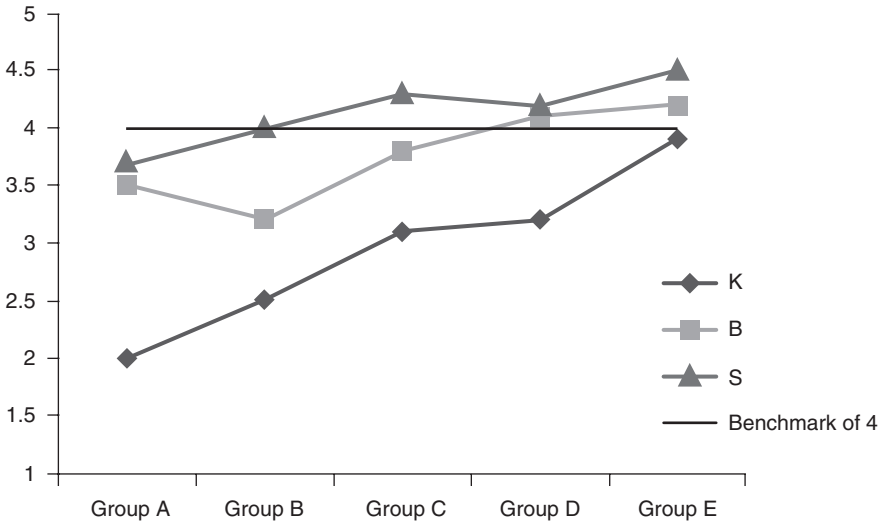

Fig. 6.3 Line graphs depicting final Knowledge, Behavior, and Status scores for the Nutrition problem for five demographically identified groups. Note that Groups B, C, D, and E attain the benchmark of 4 for Status, Groups D and E attain the benchmark for Behavior, and no groups attain the benchmark for Knowledge (Group E is close). Note differences between groups in levels of final Knowledge, Behavior, and Status scores. Note that comparisons are facilitated using the mean (averaged) scores

6.2.4 The *P-Value* in Large Dataset Research

In analyzing very large datasets, it is often the case that most statistical tests will show significance for findings that may have minimal differences. This is because of the very large numbers of cases in the sample that affect the denominator of the statistical calculation, which in the case of the t statistics is the standard error that decreases as the sample n increases. This artificial decrease in the p-value as sample size increases renders a statistical test useless at best and misleading at worst [9–10]. Therefore it is critical to provide alternative measures of significance when examining patterns in very large datasets. The same is true for smaller datasets, in which the converse argument applies: sample size may be too small to discern significance. Thus, a project outcome may be statistically significant, but not be clinically significant; or clinically but not statistically significant [10]. For these reasons, and also to understand the magnitude of change, it is important to consider the clinical or practical significance of outcomes and *p-value*s should be considered along with effect size, sample size, and study design [9–10].

6.3 Clinical or Practical Significance

Clinical significance is a health outcome change (positive or negative) that, from the perspectives of providers and/or individuals receiving care, is considered important or worthwhile, and/or leads to a change in health care management [11]. Few standards exist for determining clinical significance of outcomes, and further research is needed to develop and validate methods of quantifying and calculating clinically important changes for diverse problems and populations. Some researchers suggest that a clinically significant change may be estimated using the standard deviation or standard error of the mean (SEM) within a study (SEM is described in Chapter 5) [11]. Others suggest using effect size [12–16] as described below. Regardless of the methods that may offer guidance for statistical and clinical significance of findings, validation by clinical experts and affected individuals is key to interpretation of significance and should be a component of decision making to guide further care and research.

6.3.1 Effect Size (Clinical or Practical Significance of $P_{change} = P_{Time2} - P_{Time1}$)

When reported alongside measures of statistical significance, effect sizes such as Cohen's *d* help determine the practical meaning of results [12–14]. Effect sizes represent a collection of standardized and unstandardized indices that describe the magnitude of differences between means and the strength of associations among variables [13]. In particular, Cohen's *d* is a function of the size of the mean difference, sample size, and the correlation between the paired scores [14]. Cohen's general guidelines (1992) for interpreting effect size are as follows: small (0.2),

medium (0.5), and large (0.8) [14]. In a study evaluating the differences between statistical significance (p-values) vs clinical significance (Cohen's *d*) using KBS ratings [4], findings demonstrated differences relative to changes in outcome scores, highlight the positive relationship between statistical significance and sample size, and suggested effect size benchmarks for use in practice [16]. Compared to Cohen's d, *p*-values may exaggerate the magnitude of outcome for a large sample or mask influence of interventions for a small sample [16]. To fully report effect size results the following format is recommended: *A paired-samples t-test indicated that scores were significantly higher for females* ($M = 27.4$, $SD = 6.41$) *than for males* ($M = 18.1$, $SD = 9.33$), $t(699) = 21.3$, $p < 0.001$, $d = 0.88$ [5].

6.3.2 Interpretation of Effect Size (Clinical or Practical Significance)

State that there was a significant improvement in outcome; and then state whether the magnitude of the change was consistent with a small, medium, or large effect size. For example: *the significant improvement in Nutrition knowledge of 0.40 (d = 0.50) was consistent with a medium effect size based on Cohen* [14–16].

6.4 Associations

The PIO MM specifies relationships that explain the application of an intervention to a problem within a specified context. The purpose of testing associations statistically is to determine whether the relationships between PIO MM concepts are significant, and thus provide information to substantiate the outcomes of our interventions. In the above analyses we answered the question: *Do we make a difference?* Using correlations below, we attempt to answer the question: *What do we do that makes a difference?* However, given that in real life practice we are likely using a single group before and after design, we are describing patterns that show associations and relationships, not causation.

6.4.1 Correlation

A mutual relationship expressed as a single number between 0 and 1 that describes the degree of relationship between two variables (Fig. 6.4). To fully report Pearson's correlation results the following format is recommended: $r(XX) = .XX$, $p = .XX$, in which (XX) is the degrees of freedom [5]. Correlations may be used to evaluate relationships between PIO MM variables. Using Pearson's r (parametric) or Spearman's rho (non-parametric), analyze correlations between counts of any PIO MM variables that may be meaningfully counted (e.g. number of problems, interventions, or s/sx per person) [2–3]:

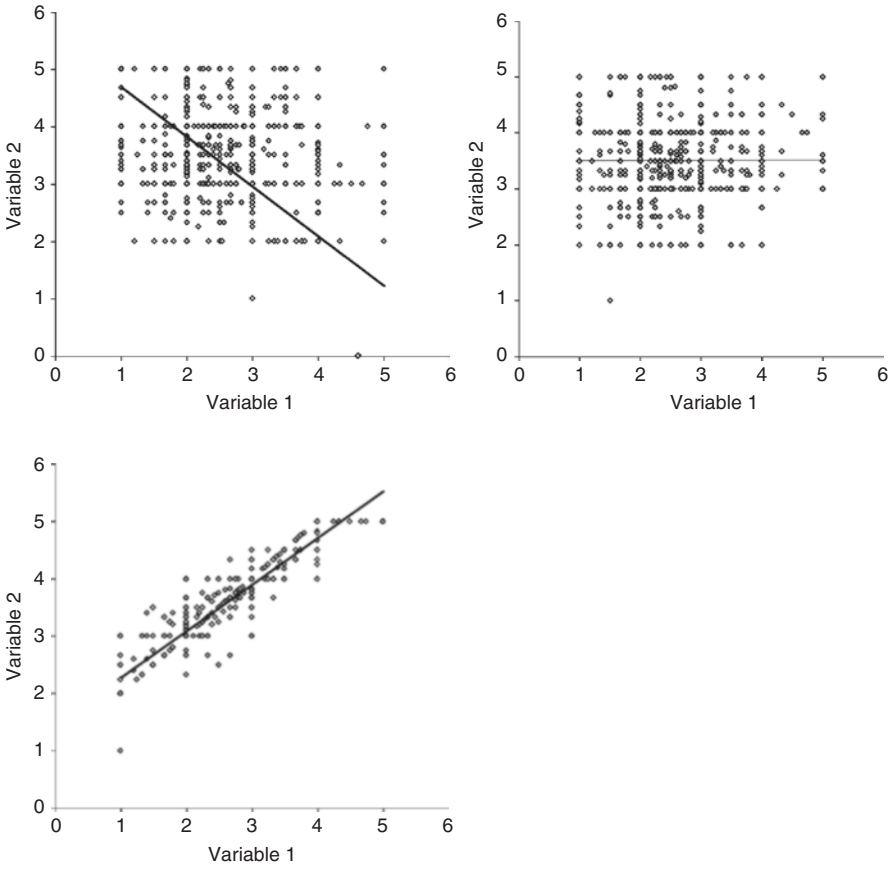

Fig. 6.4 Scatter plots with trend lines depicting correlations between variables: (**a**) Negative association ($r(744) = -0.53$, $p < 0.001$). (**b**) No relationship ($r(796) = 0.004$, n.s.). (**c**) Positive association ($r(902) = 0.932$, $p < 0.001$). If a dependent and an independent variable are compared in a scatter plot, the dependent variable is shown on the Y axis (vertical axis) and the independent variable is shown on the X axis (horizontal axis)

- Are number of problems and number of interventions correlated?
- Are number of signs/symptoms and number of interventions correlated?
- Are number of signs/symptoms and number of problems correlated?

As above, are any of the PIO MM independent variables related to final outcome attainment or change in outcome?

- Are number of problems associated with final outcome attainment or change in outcome?
- Are number of interventions associated with final outcome attainment or change in outcome?

- Are number of signs/symptoms associated with final outcome attainment or change in outcome?
- Is a risk score or other metric that has been developed from PIO MM variables associated with other independent variables? With final outcome attainment or change in outcome?

6.4.2 Regression

Regressions examine multiple correlations. Especially useful for intervention effectiveness research, quality improvement activities, and program evaluation is a hierarchical regression analysis accounting for demographics and P_{Time1} in the first steps [2–3], to examine how much of the outcome can be described by interventions (modeling by category or problem). or hierarchical regression analysis accounting for demographics and P_{Time1} in the first steps [2–3].

6.4.3 Interpretation of Correlations

Findings of correlation are reported as relationship and association (e.g. the number of signs/symptoms was positively associated with the number of interventions; increasing numbers of interventions were related to increasing outcome scores). To fully report correlation results the following format is recommended: Intervention (totals, types) were positively associated with outcomes (overall and/or by problem) r(XX) =. XX, p =. XX [5]. To fully report Spearman's rho correlations the following format is recommended: r_s = .XX, p =. XX [5]. To fully report regression results the following format is recommended: Accounting for individual characteristics (demographics and baseline assessments), the interventions explained XX% of the outcome (R^2 =. XX, F(X, XXX) = XX.XX, p =. XX).

6.4.4 Survival Analysis (P_{Time1}, P_{Time2}, … P_{TimeX})

Survival analysis is a technique used to study the length of time to the occurrence of an event. Use of Kaplan Meier curves enables estimation of time to an event (such as attaining an outcome) allowing for differences in the length of services (Fig. 6.5). The probability of the event at any point is estimated from the cumulative probability of the event for each of the preceding time intervals. The model precision depends on the number of observations [2, 17–18]. Time to outcome has been evaluated in intervention effectiveness research to examine intervention patterns by problem [16]. Considerable variation in stabilization occurred by problem [16].

Interpretation of Kaplan Meier curves. State what can be known from the findings: "Time to outcome of Nutrition problem stabilization varied significantly depending on the age cohort (p=.03), with older cohorts (ages 25–30 and 30–35) stabilizing earlier than a younger cohort (ages 13–17)."

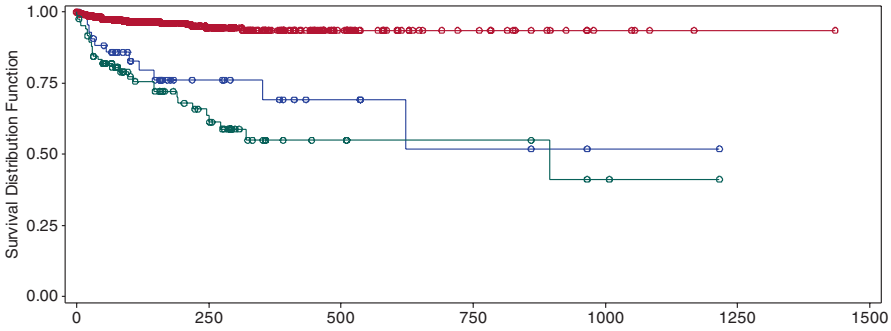

Fig. 6.5 Time to outcome showing differences between three variables using Kaplain Meier curves

6.4.5 Cross Tabs and Chi-Square (χ^2)

As described in the sample characteristics above, cross tabs may be used for identifying significant relationships among the frequencies of categorical variables using the χ^2 (Chi-Square) statistic [2]. As described above, to fully report Chi-Square analysis the following format is recommended: χ^2 (X, N = XX) = 0.XX, p =. XX [5].

6.5 Generalizability

In intervention effectiveness research, quality improvement activities, and program evaluation, the emphasis is on application of an intervention within a population of interest in a real world setting. Findings of these projects are more likely to have external validity and therefore are also more likely to be generalizable to the populations and settings similar to those in which they were conducted. In the case of retrospective studies using existing data, the findings are likely to be generalizable to populations with similar demographics and setting characteristics as the individuals in the existing data. The retrospective observational nature of such studies may limit the ability to make predictions based on results; however, accrual of a body of evidence in which numerous separate observational studies demonstrate similar findings lends confidence in generalizability beyond the study sample. The continuous improvement PDSA cycles are based on this notion, extending the units and populations involved in PDSA cycles and extending quality improvement activities as they are shown to be effective strategies in practice [20].

The inferential statistical tests discussed here will provide the foundation for understanding the effectiveness of interventions, quality improvement activities, and health improvement programs; as well as validating and extending the evidence base for health care. There are many additional techniques that can be used to relate the PIO MM variables that may be employed successfully if attention is given to the specific assumptions that relate to each test or technique. Care must always be taken

Table 6.1 Results statements for inferential analyses, with standard reporting conventions

Comparisons analyses: • Significance of outcome over time (P_{Time2} compared to P_{Time1}) overall: *There were differences in outcome after intervention overall*. • Significance of differences by group or other characteristic in outcome (final scores, change scores): *There were differences in (benchmark attainment, improvement) by (group, problem)*. • Significance of differences in sample characteristics between groups: *There were differences in (characteristic) between groups in the sample*. • Significance of intervention differences by group, problem: *There were differences in intervention (totals, types, fidelity) by (group, problem)*.
Reporting conventions for comparisons test results [5]:
Wilcoxon signed-rank test Z = X.XX, p =. XX. t-test t(XX) = X.XX, p =. XX. Chi-Square (χ^2) (X, N = XX) = 0.XX, p =. XX. Confidence intervals M = XX.X, 95% CI [XX.X, XX.X].
ANOVA or ANCOVA main effects F(1, XXX) = X.XX, p =. XX. ANOVA or ANCOVA interaction results F(2, XXX) = X.XX, p =. XX. Kruskal-Wallis H H(X) = X.XX, p = .XX [(X) is the Chi-Square value].
Cohen's d d = .XX.
Correlations analyses • Correlations (associations or relationships) between interventions and outcomes: *Intervention (totals, types) were positively associated with outcomes (overall and/or by problem)* Pearson's r(XX) =. XX, p =. XX or Spearman's rho: r_s = .XX, p =. XX • Regression of independent variables on outcome: *Accounting for individual characteristics (demographics and baseline assessments), the interventions explained XX% of the outcome* (R^2 =. XX, F(X, XXX) = XX.XX, p =. XX. For significant predictors in regression models β =. XX, p =. XX accounting for XX% of variance

to adhere to the limitations of the design, variables, data, and statistical test when interpreting findings.

Results statements based on PIO MM that report the findings of inferential statistics tests are summarized in Table 6.1.

The statistical tests described in this chapter are a few of the approaches that may be useful for evaluating comparisons and associations for significance (p-values) and magnitude (effect size measures). All results should be evaluated in relationship to the context and population by clinical experts. Further exploratory data analysis (EDA) [19] techniques may also be applied to discover patterns in either the raw data or in results, as described in the next chapter.

Reflection Questions

- How do descriptive and inferential statistics differ? How does each complement the other?
- In the ideal intervention effectiveness research or program evaluation project, what analysis would be used? If the ideal is not achievable, what analysis is supported by the variables that exist?
- In what way will the analysis approach enable evaluation of the interactions between the concepts of PIO MM?

References

1. Gay JM. Clinical epidemiology & evidence-based medicine glossary: experimental design and statistics terminology. Washington State University [Internet], 2010. [cited 2017 May 12] Available from: http://people.vetmed.wsu.edu/jmgay/courses/GlossExpDesign.htm
2. Dunn OJ, Clark VA (2009) Basic statistics: a primer for the biomedical sciences. Wiley
3. McDonald JH (2009) Handbook of biological statistics. Sparky House Publishing, Baltimore, MD
4. Engineering statistics handbook, do the observations come from a particular distribution? [Internet]. [cited 2017 May 12]. Available from: http://www.itl.nist.gov/div898/handbook/prc/section2/prc213.htm
5. American Psychological Association (APA) (2009) Publication manual of the American Psychological Association, 6th edn. American Psychological Association, Washington, DC
6. Monsen KA, Brandt JK, Brueshoff B, Chi CL, Mathiason MA, Swenson SM, Thorson DR (2017) Social determinants and health disparities associated with outcomes of women of childbearing age who receive public health nurse home visiting services. J Obstet Gynecol Neonatal Nurs 46(2):292–303
7. Camp RC, Tweet AG (1994) Benchmarking applied to health care. Jt Comm J Qual Improv 20(5):229–238
8. Monsen KA, Radosevich DM, Johnson SC, Farri O, Kerr MJ, Geppert JS (2012) Benchmark attainment by maternal and child health clients across public health nursing agencies. Public Health Nurs 29(1):11–18
9. Chatfield C (1995) Problem solving: a statistician's guide. Chapman & Hall/CRC
10. Too big to fail: large samples and the P-value problem [Internet]. [cited 2017 May 12]. Available from: http://www.galitshmueli.com/system/files/Largesample-12-6-2012.pdf 2012
11. Page P (2014) Beyond statistical significance: clinical interpretation of rehabilitation research literature. Int J Sports Phys Ther 9(5):726
12. Greenwald AG, Gonzalez R, Harris RJ, Guthrie D (1996) Effect sizes and p-values: what should be reported and what should be replicated? Psychophysiology 3(2):175–183
13. Kotrlik JW, Williams HA (2003) The incorporation of effect size in information technology, learning, and performance research. Inf Technol Learn Perform 21(1):1–7
14. Cohen J (1992) A power primer. Psychol Bull 112(1):155–159
15. Johnson KE, McMorris BJ, Raynor LA, Monsen KA (2013) What big size you have. Using effect sizes to determine the impact of public health nursing inter-ventions. Appl Clin Inform 4:434–444
16. Monsen KA, Farri O, McNaughton DB, Savik K (2011) Problem stabilization: a metric for problem improvement in home visiting clients. Appl Clin Inform 2:437–446
17. Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. J Am Stat Assoc 53:457–481
18. Freedman DA (2012) Survival analysis. Am Stat:110–119
19. Tukey JW (1977) Exploratory data analysis. Addison-Wesley, Reading, MA
20. Institute for Healthcare Improvement. The science of improvement: how to improve [Internet]. [cited 2017 May 12]. Available from: http://www.ihi.org/resources/pages/howtoimprove/scienceofimprovementhowtoimprove.aspx

Exploratory Data Analysis

7

7.1 The Development of Exploratory Data Analysis

Exploratory Data Analysis (EDA) is an approach advocated by renowned statistician J. W. Tukey and others. It uses data visualization as applied to raw data or summarized information (Chapter 5) from a dataset to understand relationships within a dataset. It may be used to discover patterns which can then be tested using standard inferential statistics (Chapter 6) [1]. Visualization is an effective method for pattern discovery because the human primary visual cortex is more efficient and effective interpreting shapes and colors than with numbers as a means to identifying patterns and distinguishing relationships [2–9]. Thus, the use of EDA helps to identify outliers, discover otherwise-hidden trends and patterns, and suggest hypotheses for statistical evaluation. It should be noted that EDA has been called a 'fishing expedition' rather than true research; however, Tukey and colleagues promoted greater use of EDA to discover testable hypotheses as a high priority to advance research [1].

In the 1970s, technology enabled machine-generated data visualization techniques, and this technology gradually became more accessible to researchers. Subsequently, more sophisticated and advanced visualization methods have been developed to extend the use of EDA to a growing scientific community through applications such as Excel, R, Tableau, and D3 [1, 10–13]. These visualization technologies together with large datasets are well suited to demonstrate the complexity of health care; particularly the complexity of multidimensional tailored interventions provided by nursing. Such complexity is difficult to discern using traditional statistical tests, as demonstrated by Frances Anscombe, whose famous Anscombe's Quartet comprises four datasets of eleven points each that have the same statistical properties, yet appear strikingly different when plotted in a graph. Anscombe's quartet demonstrates the importance of EDA as an integral part of a comprehensive data analysis [14]. Visualization is particularly important in understanding patterns in big data [1, 5–7, 14].

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_7

7.2 Interpretation of Exploratory Data Analysis

When working with EDA to examine a dataset it is often helpful to keep a list of possible patterns for further exploration and analysis. Combinations of data from the variables operationalizing the PIO MM model that may be of interest have been described in several examples. These are merely for illustration purposes to offer a starting point for identifying potential relationships within and among the variables of your project. It may be useful to create a table of variables such as the one shown in Table 7.1 in order to aid in tracking findings of EDA for PIO MM.

Upon discovery of an interesting pattern, hypotheses may be generated and tested. When findings reveal a significant pattern, it is important to validate it by review of clinical experts and by searching the scientific literature for previous documentation of the pattern. This process may lead to new research trajectories or changes in care process or programs.

7.3 Visualization Techniques

Numerous visualization techniques have been used to identify patterns in health care data. This chapter is not intended to be an exhaustive source for information or instructions regarding visualization techniques; rather, the techniques that are presented are examples that will provide a starting point for further development of visualization skills. Visualization skills, as with any new skills, improve with practice. Several iterations of a single visualization may be necessary to achieve a desired image and further iterations based on that image may be necessary to detect patterns that are potentially meaningful. For purposes of data cleaning and pattern detection there are two techniques that are simple and readily learned with typical spreadsheet technology: heat maps and multi-series line graph graphs. These are presented in more detail, followed by examples of other visualization techniques that have been used in health care research, quality improvement activities, and program evaluation.

7.3.1 Heat Map

A heat map is a matrix that is shaded to show relationships among variables arranged on the x and y axes [6]. Heat maps may be used to visualize multiple categorical variables, in order to rapidly detect a hierarchy among large multivariate datasets [6]. Heat maps developed using Excel and SPSS have been used to enable pattern detection of problems, interventions, and outcomes [4, 9, 15, 16]. Color saturations in the heat map may reveal the rank-order among different groups for relative comparisons. However, the color saturations may be specified to show magnitude by specifying the level of saturation for specific values [6]. Data tables of any type may be used as a basis for heat maps Fig. 7.1.

a

Problem	1	2	3	4	5	Grand Total
Medication regimen	9	112	47	16	9	193
Mental Health	19	73	52	17	4	165
Income	15	72	33	21	9	150
Nutrition	15	60	53	14	2	144
Pain	7	58	43	27	3	138
Substance use	19	40	39	18	5	121
Communication with community resources	13	48	23	17	6	107
Residence	7	46	23	13	5	94
Circulation	6	45	34	5	2	92
Communicable/infectious condition	11	37	23	11	8	90
Physical activity	5	44	21	5	3	78
Health care supervision	6	25	23	11	4	69
Pregnancy	8	38	10	4	1	61
Neuro-musculo-skeletal function	4	25	10	14	3	56
Respiration	10	24	14	4	2	54
Social contact	5	20	20	7	2	54
Role change	2	24	15	9	2	52
Sleep and rest patterns	3	21	17	7	1	49
Caretaking/parenting	1	15	15	12		43
Personal care	4	18	15	6		43
Urinary function	10	15	6	1	1	33
Grief	4	16	8	4		32
Skin	2	14	9	6		31
Growth and development	1	8	14	7		30
Sanitation	5	15	8	1	1	30
Cognition	7	11	5	4		27
Digestion-hydration	2	15	7	3		27

b

Problem	1	2	3	4	5	Grand Total
Medication regimen	9	112	47	16	9	193
Mental Health	19	73	52	17	4	165
Income	15	72	33	21	9	150
Nutrition	15	60	53	14	2	144
Pain	7	58	43	27	3	138
Substance use	19	40	39	18	5	121
Communication with community resources	13	48	23	17	6	107
Residence	7	46	23	13	5	94
Circulation	6	45	34	5	2	92
Communicable/infectious condition	11	37	23	11	8	90
Physical activity	5	44	21	5	3	78
Health care supervision	6	25	23	11	4	69
Pregnancy	8	38	10	4	1	61
Neuro-musculo-skeletal function	4	25	10	14	3	56
Respiration	10	24	14	4	2	54
Social contact	5	20	20	7	2	54
Role change	2	24	15	9	2	52
Sleep and rest patterns	3	21	17	7	1	49
Caretaking/parenting	1	15	15	12		43
Personal care	4	18	15	6		43
Urinary function	10	15	6	1	1	33
Grief	4	16	8	4		32
Skin	2	14	9	6		31
Growth and development	1	8	14	7		30
Sanitation	5	15	8	1	1	30
Cognition	7	11	5	4		27
Digestion-hydration	2	15	7	3		27

Fig. 7.1 Problems by score frequencies shaded with varying saturations of one color (*hue*) to reveal most frequent scores in the sample using conditional formatting functionality in Excel, with (**a**) shading for all values together, and (**b**) shading showing the relative frequencies per row

In Fig. 7.1, the number of each value of a rating scale is provided for various health problems, with more saturation for proportionately higher frequencies. Differences in shading technique within the same table [overall (a) and by row (b)] may reveal different patterns. For example, shading by row (b) reveals that 2 being the most frequent score is consistent for many problems but not all. Interesting patterns in distribution of scores by problem may lead to new questions, iterations in table development, and further exploration leading to hypothesis generation regarding particular problems. In this example, problems with higher proportional frequencies of the lowest score as detected by deeper saturation (Substance use, Respiration, Urinary function, Sanitation, and Cognition) may become the basis for forming interesting case cohort groups for further investigation.

7.3.2 Line Graph

A line graph is commonly used to display data across time but may also be used to display any continuous data for one or more variables in order to show relationships between the variables. Shown previously in Chapter 6, Figs. 6.2 and 6.3 are examples of line graphs. Line graphs link data to identify an overall shape of data and emphasize relationships among features [7]. Examples are provided for use of line graphs to discover patterns in outcome data, followed by further description of three types of line graphs: Line graph with trend line, parallel coordinates, and streamgraph.

Figures 7.2 and 7.3 are presented as examples of examining outcomes using the final score (Fig. 7.2) similar to a benchmark attainment visualization in Fig. 6.3; and using a change score (Fig. 7.3). There are two line graphs in each figure representing comparison groups within the sample (Fig. 7.2).

The line graphs in Fig. 7.2 depict three overall measures (K, B, and S averaged across all problems by individual) before and after intervention for two groups

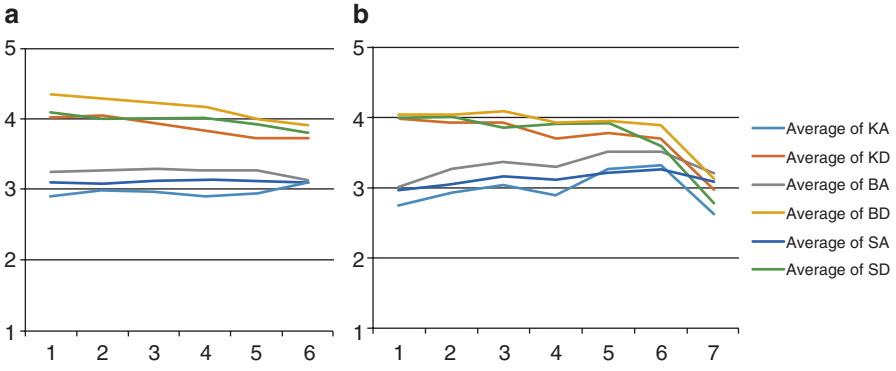

Fig. 7.2 Knowledge, behavior, and status scores for two groups (**a** and **b**) before and after intervention, displaying values by a 7-value metric, with (**a**) demonstrating consistent improvement after intervention across the metric values, and (**b**) demonstrating decreasing improvement across the metric values

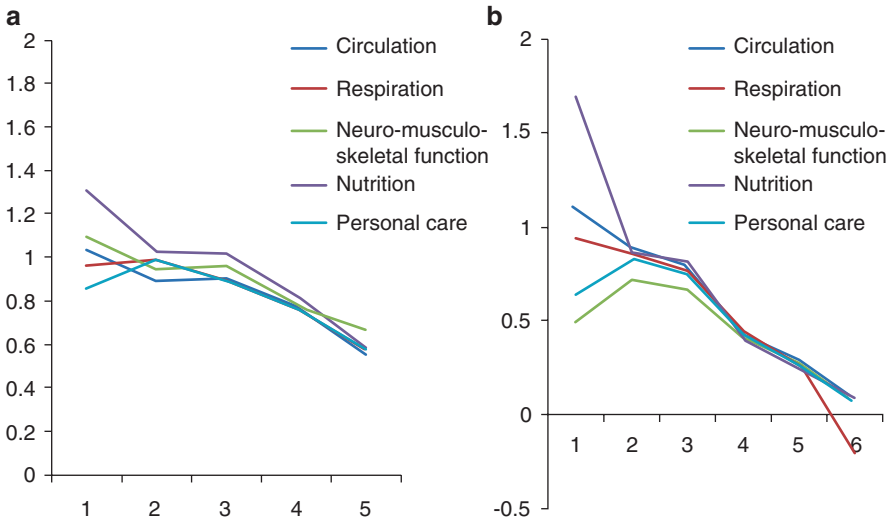

Fig. 7.3 Change in knowledge scores for five selected problems for two groups (**a** and **b**), displaying values by a 7-value metric, with (**a**) demonstrating consistent improvement after intervention across the five problems with a gradual decrease as metric values increased, and (**b**) demonstrating sharply decreasing improvement after intervention as metric values increased

stratified by a 7-value metric. The parallel lines in graph A show the consistent improvement from before (Knowledge Admission, Behavior Admission, Status Admission) to after (Knowledge Discharge, Behavior Discharge, Status Discharge) intervention for the first group. The converging and crossing lines in graph B show that as values in the metric increased for the second group, there was less improvement and finally, no improvement or worsening outcomes. In these images, both before and after KBS scores show interesting differences by group and according to the metric. These observed patterns support hypothesis testing looking at differences between groups and by metric values that may be tested using a two-way ANOVA with interaction analysis. Significant patterns may be novel or may have a basis in the literature (Fig. 7.3).

The line graphs in Fig. 7.3 depict change in Knowledge scores for five selected problems. The graphs A and B represent the sample by the same two groups and 7-value metric as in Fig. 7.2. In graph A, lines show a similar pattern of consistent improvement, gradually decreasing as the lines converge. In graph B, change scores have more variability and a steeper slope showing less and less improvement after the metric value of 3. These images of change in knowledge for these selected problems reveal possible relationships by group and metric value. Similar to the patterns discovered in Fig. 7.2, these observed patterns support hypothesis testing looking at differences between groups and by metric values that may be tested using a two-way ANOVA with interaction analysis.

Taken together the graphs in Figs. 7.2 and 7.3 provide strong support for a hypothesis that groups A and B differ in outcome attainment after intervention.

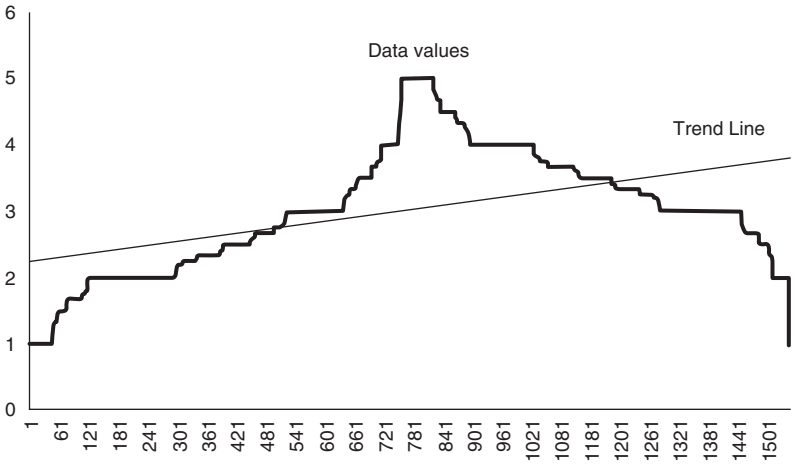

Fig. 7.4 Data displayed continuously may have an irregular shape that may mask the overall trend. This example of line graph also shows the *trend line*. This example of values (such as interventions over time in days) shows the general trend in the data shape, created in Excel using the trend line functionality

This hypothesis should be tested, with comparisons of the demographic and setting characteristics as supporting information for the report to stakeholders. Pattern discovery should be validated by clinical experts and the literature.

Line Graph with Trend Line

A line graph with trend line (Fig. 7.4) displays data values in linear fashion and includes a regression line that shows the trend in the data [4]. Algorithms to show responsiveness to interventions have been depicted using this method [17]. Line graphs are an approximate connection among data unless the algorithm and procedures for handling missing data have been provided [7]. Adding a trend line to any line graph may aid in understanding an overall pattern that may not be readily obvious (Fig. 7.4).

Parallel Coordinates

Parallel coordinates display is used to show relationships between case-specific individual lines that connect data from several categorical or numeric variables on multiple vertical axes [5, 18, 19]. Normalization of the scales represented on each of the axes based on data range enables comparison across diverse data ranges [20]. Note that horizontal positions of the axes can influence perceived associations, thus it is important to evaluate various versions of the display in which [19, 20]. Parallel coordinates displays developed using Tableau have been used to detect nurse-specific intervention effectiveness patterns within large datasets (Fig. 7.5) [9].

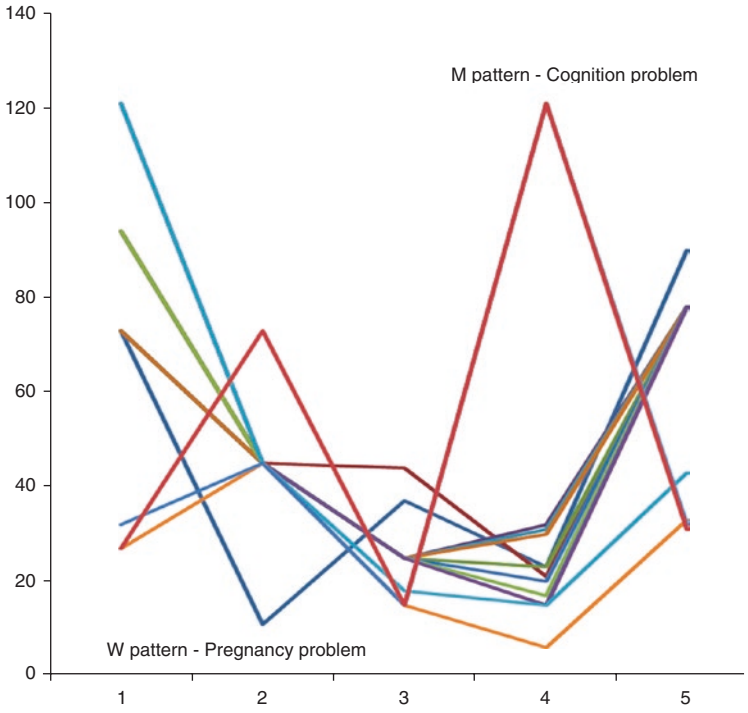

Fig. 7.5 Data that are normalized cross variables and connected to show relationships among separate attributes is a parallel coordinates graph. This example was created in Excel to show the relationships among data for five attributes across 14 problems including Cognition (*red line* with M pattern) and Pregnancy (*dark blue line* with W pattern)

Table 7.1 This table describes the PIO MM variables and provides a matrix that may be completed with notes regarding observed patterns in order to aid in tracking findings of EDA for PIO MM

Variable	Problem characteristics or metric	Intervention characteristics or metric	Population characteristics or metric	Setting characteristics	Interventionist characteristics	Outcome 1 measure or metric	Outcome 2 measure or metric	Outcome 3 measure or metric
Problem								
Intervention								
Population								
Setting								
Interventionist								
Outcome 1								
Outcome 2								
Outcome 3								

In the example provided in Fig. 7.5, values for 14 problems (the sample) are arranged on five axes, connecting data for the five separate variables in a parallel coordinates display. Colored lines are used to identify the problems in the sample. It may be observed that there are notable patterns that differ by problem and these patterns may be interesting to examine for each of the variables. Some of these may be the very high value of attribute 4 for the Cognition problem compared to all other problems, the M-shaped pattern seen with the Cognition problem, and the W-shaped pattern seen with the Pregnancy problem. Note that when lines converge it may be difficult to identify lines by their color. For this reason, it is helpful to analyze and interpret the parallel coordinates graph interactively in order to view the data simultaneously and verify the pattern with the data.

There are numerous other options for visualizing data to better understand the dataset, detect patterns, and generate hypotheses. Use of these methods consistently will increase skill levels and capacity to work with large datasets.

Reflection Questions

- What is the value and/or importance of EDA in for your project? At what point in the study is it best to employ EDA?
- Give an example of several data points that may be summarized in a single visualization, and how this may improve or decrease interpretability of the data.
- Imagine finding a novel pattern based on PIO MM variables using a heat map or a line graph. How will you determine if this pattern should be evaluated using hypothesis testing?

References

1. Tukey JW (1977) Exploratory data analysis. Addison-Wesley, Reading, MA
2. Monsen KA, Peterson JJ, Mathiason MA, Kim E, Lee S, Chi CL, Pieczkiewicz DS (2015) Data visualization techniques to showcase nursing care quality. Comput Inform Nurs 33(10):417–426
3. Dzemyda G, Kurasova O, Zilinskas J (2013) Strategies for multidimensional data visualization. In Multidimensional data visualization: Methods and applications. Springer, New York, NY, pp. 5–40
4. Lee S, Kim E, Monsen KA (2015) Public health nurse perceptions of Omaha System data visualization. Int J Med Inform 84(10):826–834
5. Kirk A (2012) Data visualization: a successful design process. Packt, Birmingham
6. Gehlenborg N, Wong B (2012) Points of view: heat maps. Nat Methods 9(3):213
7. Aigner W, Miksch S, Schumann H, Tominski C (2011) Visualization of time-oriented data. Springer, London
8. Kim E, Monsen K, Pieczkiewicz D (2013) Visualization of Omaha System data enables data-driven analysis of outcomes. In: American Medical Informatics Association Annual Meeting, Washington, DC

9. Monsen KA, Peterson JJ, Mathiason MA, Kim E, Votova B, Pieczkiewicz DS (2017) Discovering public health nurse–specific family home visiting intervention patterns using visualization techniques. West J Nurs Res 39(1):127–146
10. Mendenhall W, Beaver RJ, Beaver BM (2012) Introduction to probability and statistics. Cengage Learning, Boston, MA
11. What is R? [Internet]. [cited 12 May 2017]. Available from https://www.r-project.org/about.html
12. What is Tableau? [Internet]. [cited 12 May 2017]. Available from http://www.tableau.com/products/desktop#TGETWeLmfvacBMZt.99
13. D3 Data Driven Documents [Internet]. [cited 12 May 2017]. Available from https://d3js.org/
14. Anscombe FJ (1973) Graphs in statistical analysis. Am Stat 27(1):17–21. doi:10.2307/2682899
15. Green SB, Salkind NJ (2010) Using SPSS for windows and macintosh: analyzing and understanding data. Prentice Hall, Upper Saddle River, NJ
16. Radhakrishnan K, Monsen KA, Bae SH, Zhang W (2016) Visual analytics for pattern discovery in home care. Appl Clin Inform 7(3):711–730
17. Chi C (2015) Data mining for translation to practice. Presented at 2nd international conference on research methods for standard terminologies, West Paul, MN
18. Harter JM, Wu X, Alabi OS, Phadke M, Pinto L, Dougherty D, Petersen H, Bass S, Taylor II RM (2012) Increasing the perceptual salience of relationships in parallel coordinate plots. In: IS&T/SPIE Electronic Imaging, 2012 Jan 22. International Society for Optics and Photonics, pp 82940T–82940T
19. Inselberg A (2009) Parallel coordinates: visual multidimensional geometry and its applications. Springer, New York, pp 1–554
20. Few S. Multivariate analysis using parallel coordinates. Perceptual Edge web site. [Internet]. [cited 12 May 2017]. Available from http://www.perceptualedge.com/articles/b-eye/parallel_coordinates.pdf. Published 12 Sept 2006. Updated 2015

Ethical Considerations

8

8.1 Minimal Risk

We discuss *minimal risk* [1] because it underlies the notion that intervention effectiveness research, quality improvement activities, and program evaluation projects should be conducted in alignment with the definition of minimal risk. Further, it is not possible nor ethical to implement research controls such as randomization in which people receive effective vs. risky treatments or interventions of unknown effectiveness in response to healthcare needs. The ethical considerations for intervention effectiveness research, quality improvement activities, and program evaluation have commonalities and also differences based on setting and use of existing or new data. In determining the level of risk for intervention effectiveness research, quality improvement activities, and program evaluation, consider the intervention or interventions within a program as they align with the definition of minimal risk [1], and whether the findings of your project generate new, generalizable knowledge [1–6].

According to United States federal regulations governing research with humans (the Common Rule) [1], minimal risk is defined as "the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests" (also known as the "daily life standard") [1, 7]. Intervention effectiveness research studies align with minimal risk in that such studies test the effectiveness of a *safe, efficacious or effective* intervention. Quality improvement projects align with minimal risk in that *evidence-based interventions*, variations in procedures used to implement *effective interventions*, or observed improvements *on routine care* are typically evaluated. Program evaluation projects align with minimum risk in that evaluation of programs that are known to be *safe and effective*. Furthermore, the data used in these studies and projects should be obtained from existing databases or measures that also align with the Common Rule definition of *routine physical or psychological examinations or tests* [1]. For example, existing data from EHRs align with the Common

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_8

Rule definition because these data were generated during the routine documentation of healthcare by the healthcare clinician.

8.2 Institutional Review

When planning to conduct intervention effectiveness research, quality improvement activities, and program evaluation there is an expectation that your project should be approved by all related institutions, and should not involve unnecessary risks. Documentation of approvals from the affiliated institutions and settings is an essential first step that must be completed before initiating the project. In considering the Institutional Review Board (IRB) process relative to your project, it is important to distinguish between research and non-research relative to human subjects participation, and whether the findings of your project generate new, generalizable knowledge [2–7].

The United States National Institutes of Health National Cancer provides a definition of an Institutional Review Board: "A group of scientists, doctors, clergy, and patient advocates that reviews and approves the detailed plan for every clinical trial. Institutional Review Boards are meant to protect the people who take part in a clinical trial. They check to see that the trial is well designed, legal, ethical, does not involve unneeded risks, and includes a safety plan for patients. There is an Institutional Review Board at every health care facility that does clinical research. Also called IRB." [8].

8.2.1 Where and How to Access an IRB

While there is typically an IRB at every health care facility that does clinical research, clinical settings that do not customarily engage in research may not have an IRB. For example, small hospitals, clinics, and health departments rarely have IRBs within their organizations. When conducting intervention effectiveness research, quality improvement activities, or program evaluation in a setting that does not have its own IRB, the permission of the organization's director should be obtained and documented. In such cases it is also advisable to access the IRB of an educational institution such as a university at which a student or faculty involved in the project may be affiliated.

Each IRB has policies and procedures for reviewing proposed studies that are unique to the organization or university. It is advisable to obtain and study these policies and procedures in advance to understand the expectations related to your project. Many IRBs post their policies and procedures on-line.

8.2.2 When a Project May Be Exempt from IRB Review

Studies using existing data without variables that may identify individuals may be deemed exempt from IRB review according to US HHS policy: "§46.101 (4) Research involving the collection or study of existing data, documents, records,

pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects." [5] Special IRB procedures for such exempt studies may accessed from the IRB, and forms may be completed that describe how the data were generated, what the data contain, and whether or not there are links to identifiers. Confirmation that the study is deemed exempt from IRB review through these procedures is sufficient for documenting adherence to ethical standards. This is a necessary requirement for the results of the project to be considered for publication in the scientific literature.

8.2.3 The Special Case of Quality Improvement

Quality improvement projects may be exempt from IRB review by definition. Some university, health system, and agency IRBs have protocols for quality improvement reviews that exempt quality improvement projects from review by definition. However, it is essential to secure an official statement that your project was deemed to be exempt from IRB review in order to share findings in the scientific literature. Distinguishing quality improvement from human subjects research is essential. There are usually special IRB procedures for quality improvement studies that may be followed using forms that describe the nature of the project. Health systems should have a process in place to review quality improvement proposals to ensure that that they meet the criteria for minimal risk, usual care, and privacy and security protections identifiable data [9].

The delineation as a separate circumstance from research in ethical review for quality improvement is grounded in moral deliberation by experts such as a group from the Hasting Center, which defined quality improvement as "systematic, data-guided activities designed to bring about immediate improvements in health delivery in particular settings." [2] This is different from formal research, defined as "a systematic investigation, including research development, testing, and evaluation, designed to develop or contribute to generalizable knowledge" [9] in aspects of immediacy and generalizability [2, 9]. Scholars and clinicians agree that health care providers have a fundamental obligation to incorporate quality improvement within health care settings to systematically evaluate and learn from experience [9]. For example, quality improvement PDSA cycles may be based on practical problem solving, evidence-based interventions, review of performance data, and/or applications of theory-driven processes designed for system change [2, 10–13]. Your project may be deemed quality improvement if it will be conducted using a PDSA cycle method; is to be based on experiences and insights to identify promising improvements; will implement changes on a small scale in a real life setting; will monitor and interpret effects of the improvements; and is intended to aid in deciding whether there should be wider implementation of the improvements. Quality improvement studies may compare data from one unit or across multiple units or organizations [2, 9–13].

Further, while quality improvement may be seen as a special case in human subjects research, scholars and clinicians note important qualities that distinguish quality improvement from classical human subjects research from the perspective of

ethical review in purpose, starting/end point, design, risks, benefits, and participant obligation [9]:

- *Purpose.* In quality improvement the purpose of the PDSA cycle is to assess a process or program after implementation of an improvement in a particular setting. The purpose of research is to generate new knowledge regarding the effectiveness of an intervention that may be generalizable across settings.
- *Starting/ending point*. In quality improvement PDSA cycles are continuous and ongoing within the health care system. This means that many processes may be simultaneously undergoing quality improvement study, and these may begin and end at various points in time. In research, there is a starting point at which the plan for testing an hypothesis or research question is initiated according to the research plan. Likewise, research ends when the hypothesis is supported (or not supported) and/or the research question is answered.
- *Risks.* In quality improvement, PDSA cycles do not increase risk to patients and patients are not consented. In research, patients must be consented, and any risks must be explained during the consent process.
- *Benefits.* In quality improvement and in research, benefits may or may not be experienced by the participants; however it is more likely that patients may benefit if a rapid PDSA environment in which quality improvements may be incorporated within the health system during the PDSA process.
- *Participant obligation.* In research, participants are under no obligation participate and may stop participating at any time. However, in quality improvement, the PDSA cycle exists within the health care context and participants are patients who receive care that are necessarily involved in the improvement process. Some scholars and clinicians assert that it is the responsibility of patients to participate as component of receiving care [9].

Criteria for IRB review of quality improvement may differ across settings. However, generally speaking your project should undergo ethical review through an IRB if it seeks to evaluate the efficacy of new or controversial interventions or treatments, rather than to assess the implementation of known interventions; when participants are randomized to receive different interventions in order to ensure confidence in the findings; when external funding from a research institute or drug company was received to support your project; and/or when interventions place participants at risk. Such studies are beyond the scope of intervention effectiveness research, quality improvement activities, and program evaluation; and further guidance should be consulted [9].

8.2.4 Minimal Risk and IRB Review

From the IRB perspective, decisions about whether or not a study involves minimal risk are critical procedural and substantive determinations [1]. Minimal risk is a necessary condition for deciding whether a study may qualify for expedited IRB

ethical review and whether a waiver of the requirement for informed consent might be permissible. In addition, the determination that a study involves minimal risk influences the way investigators describe it to prospective research participants in consent documents and during the informed consent process [1]. Judgments about minimal risk are challenging if the procedures that study participants undergo and the risks they face will be similar to ordinary care, because the risks of the ordinary care may be seen as either high- or low-risk [1].

If the IRB determines that the incremental risks of study participation, over and above receipt of usual therapy outside the study, do not exceed the daily life standard, then it can justifiably conclude that the trial involves only minimal risk [1]. The same standard is applicable when randomization occurs at the group rather than the individual level, such as in cluster-randomized trials or system-wide quality improvement tests [1]. Further, the U.S. Department of Health and Human Services (DHHS) Office for Human Research Protections (OHRP) provides guidance on study eligibility for expedited rather than full-board IRB review: the study must be judged to involve no more than minimal risk, and all study procedures must fit within one or more of the categories on a list published by OHRP/DHHS [7].

8.2.5 The Special Case of Program Evaluation

Program evaluation is part of the day to day evaluation of program effectiveness in public health and other settings in which interventions have been developed and are provided to identified populations [5, 6]. Similar to quality improvement, program evaluation takes place in the real world within existing programs and may or may not be considered research [5, 6]. The delineation of program evaluation as a separate circumstance from research has been described in a United States Centers for Disease Control and Prevention policy, according to the evaluation purpose. Program evaluation is not research when the purpose of the evaluation is to assess the success of an established program in achieving its objectives in a specific population, as described in this book. In program evaluation that is not research, the findings of the evaluation will be used to provide feedback about the program to the program in order to monitor success and/or improve the program [5, 6]. Such evaluation activity is frequently an ongoing process that is embedded within the program activities with the goal of providing information learned for immediate benefit to program and/or the individuals who participate in the program. Furthermore, the interventions that are evaluated are not novel or experimental; rather they are known by the scientific community to be effective [5, 6].

In comparison, program evaluation is considered research when the purpose is to test a new, modified, or previously untested intervention, service, or program to determine whether it is effective compared to standard interventions [5, 6]. This type of program evaluation generates new knowledge that is applicable beyond a single program, and to contribute to the knowledge base regarding program effectiveness that may be applied to other programs, settings, or populations. Thus, the purpose of program evaluation research is to generate new knowledge or contribute

to the knowledge in the scientific literature [5, 6]. Program evaluations that are considered research may not need to undergo IRB review. To determine if IRB review is required, additional criteria must be assessed, including whether the research involves human participants; and if so, whether the proposed research meet the criteria for exemption from 45 CFR part 46 as described above [5, 6].

Determining that intervention effectiveness research, quality improvement activities, and program evaluation projects satisfy the minimal-risk standard and requirements for exemption from review reduces the administrative burden of IRB review without compromising the protections afforded to participants. When making a distinction between exempt studies, quality improvement activities, and program evaluations vs. research, it is clear that the risks of the non-research interventions would not exceed the daily life standard specified in the Common Rule [1, 7]. Exemption from review of these studies by the IRB thereby adheres to the ethical mandate to ensure participant protection against exposure to excess research-related risk [1].

8.3 Informed Consent

Informed Consent Processes in the Context of Minimal Risk Research [14]. Issues regarding informed consent may apply to proposed intervention effectiveness research, quality improvement activities, and program evaluation projects using existing data and/or new data collected prospectively with participants. Institutional Review Boards and data privacy officials may provide guidance on informed consent procedures commonly used in project settings.

8.3.1 What Is Informed Consent?

The informed consent process has been defined as "an agreement obtained from a subject, or from his authorized representative, to the subject's participation in an activity" [15, p. 283] when "the person involved has legal capacity to give consent, is situated as to be able to exercise free power of choice, and is provided with a fair explanation of all material information [relevant to the choice]" [15, p. 282].

There are special considerations regarding informed consent related to minimal risk research for which a simplified informed consent process would be appropriate [14, 15]. Much of the research that is minimal risk involves no procedures for which written consent is required outside the research context; therefore an oral vs. written consent may be warranted. For example, participants in a program may give oral consent to use of their data in outcome research. When an IRB determines that a written form is appropriate, IRBs and researchers should strive for simplicity. Rather than relying solely on consent forms, investigators, institutions and IRBs should create other techniques/mechanisms for ensuring compliance and facilitating auditing functions. For example, a researcher may provide a signed and dated attestation affirming that consent was obtained; or a data field on an internet survey may be used to indicate that a subject affirmed their intent to participate before

beginning. Use of alternative models to record consent may require the IRB to approve a waiver of documentation [14, 15].

The US DHHS Secretary's Advisory Committee on Human Research Protections (SACHRP) proposes the following guiding principles for informed consent for minimal risk research that are summarized as follows [14]:

- Be concise. Include only the information about the research that a reasonable person would want to know.
- Match the informed consent process and supplemental materials to the research.
- An oral consent process should be used when it will enhance the quality of the consent process.
- Use language understandable to the study population.
- Use separate supplemental documents to convey information that is not directly relevant to the consent process.
- Risks/burdens that are immaterial or obvious need not be explicitly addressed.
- There is also no need to state the absence of risk where none exists.
- Use a simple statement regarding confidentiality and its limitations.

8.3.2 Informed Consent Processes in the Context of Existing Data

When existing data are used in de-identified form, permission to use the data for evaluation or research should have been documented at the time of data collection, and the consent process for original data capture should be documented [16]. This information may be included in an application for exempt status or exemption from review should be submitted to the IRB, describing the use of de-identified existing data. The IRB determination of exemption from review is based on the notion that analysis of de-identified data does not pose an informational risk [17].

8.4 Data Privacy and Security

Data privacy and security are of paramount importance for all research, quality improvement, and program evaluation projects. Development of a data management plan is an essential first step before data are collected or obtained.

According to the United States National Institutes of Health, *privacy* is "concerned with the collection, storage, and use of personal information, and examines whether data can be collected in the first place, as well as the justifications, if any, under which data collected for one purpose can be used for another (secondary) purpose. An important issue in privacy analysis is whether the individual has authorized particular uses of his or her personal information" [3]. Transformation of identifiers within a dataset, or pseudonymization of the data, is a method that may be used to maintain critical linkages among data records without revealing true identities of the individuals [3].

Security is "the procedural and technical measures required (a) to prevent unauthorized access, modification, use, and dissemination of data stored or processed in a computer system, (b) to prevent any deliberate denial of service, and (c) to protect the system in its entirety from physical harm" [4]. Measures to must be taken protect the data for intervention effectiveness research, quality improvement activities, and program evaluation to ensure that no breaches in security will occur [3]. Such measures may include appointing a security officer to assess and address data protection needs; use encryption to protect data during transfers; assign experts in data security to IRBs; and implement multiple layers of security such as duo factor sign-in to access data.

Institutions and agencies must adhere to statutory regulations regarding data practices, such as the Health Information Portability and Accountability Act [3]. Data practices statutes may vary by jurisdiction; therefore it is critical to be aware of the statutes that apply to your project. Institutional Review Boards and university libraries are potential resources for information about local data practices and research data management. Data shelters and other protections may be available within the organization or health system that can be used to ensure data privacy and security. Data sharing procedures such as secure file transfer links are preferable to other methods. Security may be jeopardized if data are transferred through normal e-mail messaging or file sharing solutions that are available to the general public. Use of any datasets in a laptop or other portable device that may be lost or stolen constitutes a data security risk [3].

Newly emerging software and biomedical technologies may make original de-identification and data security protections obsolete as powerful new techniques may create new identifiable linkages and reveal identities through the presence or absence of unique data patterns [17]. To ensure responsible future access and use of data it is recommended "that all investigators who will have access to data in the future will be bound by the best practices in data and confidentiality protections at the time of data collection and [will be bound by] new protections as they emerge" [17, p. 12].

When in doubt regarding minimal risk, ethical review, consent, and/or data management, it is important to consult with the IRB and/or agency director(s) about any questions regarding risk to participants, either as recipients of care, or questions related to data practices and security.

Reflection Questions

- Describe how your project meets the definition of minimal risk. Describe how the measures planned for your project meet the definition of minimal risk.
- Describe the evidence-based practice in the project. How does this pose minimal risk to research participants?
- What are the differences in IRB procedures at your institution for quality improvement or program evaluation vs. research?
- How will you ensure that there are no breaches in privacy or security after data collection?

References

1. Office for Human Research Protections [US DHHS]. Federal policy for the protection of human subjects ('Common Rule').[Internet]. [cited 12 May 2017]. Available from https://www.hhs.gov/ohrp/regulations-and-policy/regulations/common-rule/
2. Lynn J, Baily MA, Bottrell M, Jennings B, Levine RJ, Davidoff F, Casarett D, Corrigan J, Fox E, Wynia MK, Agich GJ (2007) The ethics of using quality improvement methods in health care. Ann Intern Med 146(9):666–673
3. Gostin LO, Levit LA, Nass SJ (eds) (2009) Beyond the HIPAA privacy rule: enhancing privacy, improving health through research. National Academies Press, Washington, DC
4. Turn R, Ware WH (1976) Privacy and security issues in information systems. Rand Corporation, Santa Monica, CA
5. Office for Human Research Protections. [US DHHS] Health and human services subpart A – basic HHS policy for protection of human research subjects. [Internet]. [cited 12 May 2017]. Available from https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html#46.101
6. Centers for Disease Control and Prevention [US]. Distinguishing public health research and public health nonresearch [Internet]. [cited 12 May 2017]. Available from https://www.cdc.gov/od/science/integrity/docs/cdc-policy-distinguishing-public-health-research-nonresearch.pdf
7. Joffe S, Wertheimer A (2013) Determining minimal risk for comparative effectiveness research. IRB 36(3):16–18
8. Cancer.gov. [US] Publications: definition of IRB [Internet]. [cited 12 May 2017]. Available from https://www.cancer.gov/publications/dictionaries/cancer-terms?cdrid=44679
9. Children's Hospital of Philadelphia Institutional Review Board. Quality improvement vs. research [Internet]. [cited 12 May 2017]. Available from https://irb.research.chop.edu/quality-improvement-vs-research
10. A look at Six Sigma's increasing role in improving healthcare.[Internet]. [cited 12 May 2017]. Available from https://www.villanovau.com/resources/six-sigma/six-sigma-improving-healthcare/#.WPIJ4vnythE
11. What is Six Sigma? [Internet]. [cited 12 May 2017]. Available from https://www.villanovau.com/resources/six-sigma/what-is-six-sigma/#.WPIWKfnythE
12. NHS Institute for Innovation and Improvement. The handbook of quality and service improvement tools. [Internet]. [cited 12 May 2017]. Available from http://www.miltonkeynesccg.nhs.uk/resources/uploads/files/NHS%20III%20Handbook%20serviceimprove.pdf
13. New to lean Six Sigma [Internet]. [cited 12 May 2017]. Available from https://www.isixsigma.com/new-lean-six-sigma/
14. Office for Human Research Protections [US DHHS]. Attachment A: recommended guidance on minimal risk research and informed consent. [Internet]. [cited 12 May 2017]. Available from https://www.hhs.gov/ohrp/sachrp-committee/recommendations/2015-september-28-attachment-a/index.html
15. Faden RR, Beauchamp TL (1986) A history and theory of informed consent. Oxford University Press, New York
16. Federal Register–9 Nov 1998 (Volume 63, Number 216) [Internet]. [cited 12 May 2017]. Available from https://www.hhs.gov/ohrp/news/federal-register-notices/federal-register-11-09-1998-vol-63-no-216/index.html
17. Fisher CB (2013) Decoding the ethics code: a practical guide for psychologists, 3rd edn. Sage, Thousand Oaks

Part II

Practical Guide for Using the Problem-Intervention-Outcome Meta-Model

Use the Worksheets and PIO MM Figure

9

9.1 Review of Part I

ter 1 provided an overview of intervention effectiveness research, quality improvement, and program evaluation. Though these perspectives are different, they share key similarities. At the heart of all of them is the need to know if an efficacious or research-based intervention is effective in a real-world setting, producing the intended and desired outcomes.

Chapter 2 introduced the Problem-Intervention-Outcome Meta-Model (PIO MM). This model offers a theoretical foundation for any project that aims to demonstrate intervention effectiveness, outcomes for a particular population of interest within a given context, or both.

Chapter 3 gave examples of research designs that may be constructed based on the PIO MM Meta-Model. Such designs can build on any aspect of the model to relate multiple populations of interest, problems, and interventions along with related contextual factors.

In Chapter 4, checklists for obtaining new or existing data were provided for two situations: collecting new data prospectively, and reusing existing data. Both situations call for a standardized terminology, and the Omaha System is an exemplar in this area.

Chapter 5 presented a step-by-step process for data analysis. This begins with data cleaning and preprocessing, which may include descriptive analyses of all variables. Once the data are clean and final descriptive statistics are completed, the inferential statistics discussed in Chapter 6 may be leveraged to better understand the practical and statistical significance of the data. Exploratory data analysis, as explained in Chapter 7, may reveal hidden patterns in the data for subsequent formal hypothesis testing.

Finally, Chapter 8 summarized ethical issues that are essential to consider in all research and evaluation. These include attention to minimal risk, institutional review, informed consent, and data privacy and security.

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_9

9.2 Overview of Part II

In Part II of this book we make the transition from theory to practice—that is, to designing and conducting an original project and disseminating the findings. Our essential strategy is to complete a series of worksheets that mark the milestones in such a project:

- Specify the Concepts (PIO MM Diagram)
- Know the Literature (Worksheet A)
- Define the Problem (Worksheet B)
- Describe the Intervention (Worksheet C)
- Define the Outcome (Worksheet D)
- Plan the Analysis (Worksheet E)
- Interpret the Results (Worksheet F)

These steps and their accompanying worksheets are based on the worldviews described in Chapter 1 and the PIO MM concepts (Chapter 2); and are informed by IOM guidelines for intervention efficacy and effectiveness research [1]. Blank copies of the worksheets are included in this chapter. Also included is a blank diagram of the PIO MM, which offers a simple and yet effective way to understand the core elements of the project.

Begin your study of Part II by reviewing the blank worksheets. Then, before beginning to use the worksheets, read through the remaining chapters in Part II. These give step-by-step instructions for completing each worksheet and examples of completed worksheets. At this point you will have the whole process in mind as you begin your project.

9.2.1 Examples of Projects

Note that Chapters 10–16 include examples of projects that focus on the following topics, approaches, and questions.

In all chapters, examples for three projects include:

- For intervention effectiveness research: *Was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions*?
- For quality improvement: *Was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?*

- For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?*

In selected chapters, additional examples are provided, together with completed Worksheets:

- For intervention effectiveness research: *Were there differences in health outcomes of older adults receiving home care interventions related to frailty and social determinants?*
- For intervention effectiveness research: *Were there differences in public health nurse home visiting interventions associated with improved outcomes adult and adolescent mothers with and without the mental health problem?*
- For program evaluation: *Was there a difference in health outcomes of individuals receiving 1-2 maternal-child public health nursing home visits overall? And compared to individuals receiving 3 or more visits?*

These examples and their related details—while representative of what you will find in the scientific literature—are fictitious. They were invented specifically to illustrate the concepts presented in each chapter and to show how the worksheets can be completed. Inconsistencies in the reporting of the fictitious examples have been introduced in order to provide the opportunity to critique the examples and avoid the temptation of simply copying any given example. There is one example that is not fictitious: In Chapter 10, a published article entitled *Care coordination from a strengths perspective: A practice-based evidence evaluation of evidence-based practice* [2] provides an example of data abstraction for the literature matrix in Worksheet A.

9.3 Starting the Process

9.3.1 Worksheet Review

The step by step process described in Part II is both iterative and sequential. It begins now as you review the PIO MM diagram and each of the worksheets. During your review, add notes to each worksheet related to your proposed project. These notes are placeholders for your thoughts that will be refined throughout the planning process based on scientific literature and feedback from stakeholders.

Worksheet A. PIO MM Literature Matrix Template

The PIO MM Literature Matrix Template is based on the work of Judith Garrard [3]. Use of the Matrix is described in Chapter 10

Article	Population of Interest	Problem addressed	Measure(s) of Outcome (Benchmark $P_{Time2;}$ P_{Time2}-P_{Time1})	Intervention(s) used	Measures of Intervention (including timing)	Measure of Intervention Fidelity	Demographic characteristics—Sample	Contextual factors—Setting	Contextual factors—interventionist	Analysis Methods	Comments	Complete reference
A												
B												
C												

Worksheet B. Definition of the Problem with Statement of the Gap in Knowledge

Problem

Definition of the Problem	
Population of Interest	
Background literature describing the problem (% or number of individuals in the total population)	
Background (costs incurred addressing the problem)	
Background (years of potential life lost due to the problem)	
Background (health system-related gaps in addressing the problem)	
Problem measurement instrument/scale	
Anticipated outcome and rationale	
What is not known?	

Worksheet B is based on the ideas discussed in Chapter 2, and the literature found when completing Worksheet A. Completion of Worksheet B is described in Chapter 11.

Worksheet C. Describe the Intervention Used to Address the Problem

Definition of the Intervention	
Percentage of persons who improved after intervention	
Meta-analysis showing levels of effectiveness across studies	
For whom	
Under what conditions	
Theory of causal mechanism	
Essential (core) components	
Intervention content	
Intervention adherence and involvement	
Intervention measure—amount	
Intervention measure—type	
Intervention measure—fidelity	
Intervention measure—quality	
Interventionist—qualifications	
Interventionist—training	
Interventionist—demographics	
Interventionist—organization	
Source of data	

Worksheet C is based on the ideas discussed in Chapter 2, the literature found when completing Worksheet A, and the specific measures that operationalize the intervention concepts. Completion of Worksheet C is described in Chapter 12.

Worksheet D. Define the Outcome and Related Measures or Scales

Definition of the Problem	
Measure/scale that operationalize the problem	
Use of measure in previous research	
Psychometric properties	
Validity (construct)	
Reliability (internal consistency)	
Reliability (test-retest)	
Reliability (across raters)	
Data collection—strategy	
Data collection—training	
Data collection—timing	
Source of data	

Worksheet D is based on specific measures that operationalize the concepts described in Worksheet B. Completion of Worksheet D is described in Chapter 13.

Worksheet E. Plan the Analysis Methods

Exploratory data analysis: Sample characteristics	
Exploratory data analysis: Interventions	
Exploratory data analysis: Outcomes	
Descriptive analysis: Sample	
Descriptive analysis: Interventions	
Descriptive analysis: Problems	
Descriptive analysis: Outcome P_{Time1}, P_{Time2}	
Inferential analysis: Outcome P_{Time1}, P_{Time2}	
Inferential analysis: Effect size	
Inferential analysis: Benchmark attainment	
Inferential analysis: Survival	
Time analysis	
Inferential analysis: Correlations	
Qualitative analysis	

Worksheet E is based on the information in Chapters 3–7. Completion of Worksheet C is described in Chapter 14.

Worksheet F. Results Statements and Alignment with the Literature

Sample characteristics	
How sample aligns with literature or demographics of the population	
Interventions	
How interventions align with the literature	
Outcomes	
How Outcomes align with the literature	
Relationship between interventions and outcomes	
How Intervention/Outcome Relationships align with the literature	

Worksheet E is based on the results and their alignment with the literature. Completion of Worksheet F is described in Chapter 15.

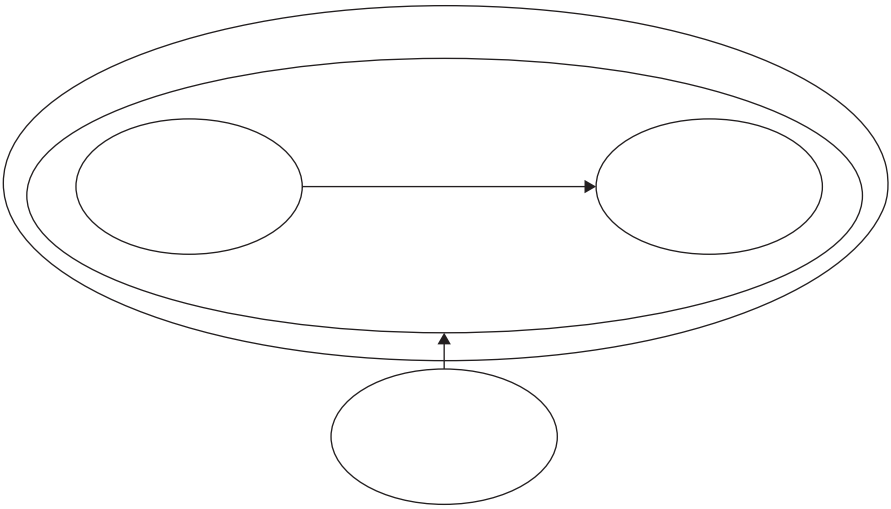

Fig. 9.1 The Project-Intervention-Outcome Meta-Model (PIO MM)

9.3.2 Complete the PIO MM Diagram

Then complete the PIO MM diagram (Fig. 9.1) to describe how the PIO MM will look in your project. Three completed examples of PIO MM diagrams are provided in Chapters 10–13. Practice completing a PIO MM diagram for your project as a starting point for your work.

After completing the PIO MM diagram and adding preliminary notes to the worksheets you are ready to begin the in-depth planning process for your proposed project.

References

1. Olsen L, McGinnis JM (2010) Redesigning the clinical effectiveness research paradigm: innovation and practice-based approaches: workshop summary. National Academies Press, Washington, DC
2. Monsen KA, Vanderboom CE, Olson KS, Larson ME, Holland DE (2017) Care coordination from a strengths perspective: a practice-based evidence evaluation of evidence-based practice. Res Theory Nurs Pract 31(1):39–55
3. Gerrard J (2010) Health sciences literature review made easy: the matrix method, 3rd edn. Jones & Bartlett Learning, Burlington, MA

Know the Literature (Worksheet A) 10

10.1 Preparing to Complete Worksheet A

It is typical to begin an intervention effectiveness research, quality improvement activities, or program evaluation study with an idea that doing things differently may improve practice or population health. You may have thought about this in some depth, and you may have a fairly well-formed plan in mind. However, to begin planning the study with a solid foundation, rationale from the scientific literature must be provided for addressing the identified problem using the proposed intervention. A thorough understanding of the current literature is essential preparation for study planning. Developing a literature matrix for the study/evaluation based on the PIO MM concepts is a convenient and rigorous way to demonstrate the rationale for conducting the study/evaluation and for substantiating the evidence for the intervention [1]. The prototype literature matrix (Worksheet A) is shown in Chapter 9, and is available on-line. It consists of rows for each article, and columns for the PIO MM concepts: Population of Interest, Problem addressed, Measure(s) of Outcome, (Benchmark P_{Time2}; P_{Time2}-P_{Time1}), Intervention(s) used, Measures of Intervention (including timing), Measure of Interventionist, Intervention Fidelity, Demographic characteristics—Sample, Contextual factors—Setting, Contextual factors—interventionist, and Analysis Methods.

The PIO MM literature matrix should be used to capture salient details regarding several studies that support or explain problem- and/or intervention-specific reports of previous research related to a population of interest. Using the matrix will enable rapid synthesis and retrieval of relevant information for planning and dissemination of PIO MM guided studies. For example, after identifying numerous articles through your search strategies and abstracting the data from the articles, you may observe that several articles describe using a similar intervention description or a particular outcome measure that may be useful in your proposed project.

Use of Excel or other electronic spreadsheets for matrix development is ideal, as the size of the matrix will be larger than standard paper formats [1]. Furthermore, the ability to sort and classify information that is recorded within a spreadsheet is

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_10

likely to enhance analysis and reveal patterns in the information that would not be easily discovered using a word processor or paper. The Chapters in Part II define these concepts in depth and provide examples of completed abstraction from research articles. This matrix layout depicts the horizontal arrangement of the concepts with each article having its own row for ease in sorting and synthesis of the content. Upon completing Worksheet A with relevant literature, the extent to which there is evidence to support your project may be synthesized and summarized.

10.2 Step-by Step Instructions for Completing Worksheet A

Standard literature search methods should be used to identify information that is relevant to the study/evaluation. Theory-guided literature search builds on standard literature search methods. Similar to all literature searches, you will seek to incorporate literature from respected sources such as reports of high quality research and highly rated evidence. In contrast to other searches, using the PIO MM as the basis for your literature search provides a theory-guided perspective for the elucidation of the problem, intervention, and outcome concepts that underlie intervention effectiveness research, quality improvement activities, and program evaluation. This will allow you to document all PIO MM concepts within the matrix so that you have sufficient body of evidence for each concept to guide the project.

Included here is an abstract of a published article with relevance for the background section for an intervention effectiveness research project entitled *A practice-based evidence evaluation of evidence-based practice* (abstract and further excerpts used with permission) [2]. We provide examples from the three projects described in Chapter 9: For intervention effectiveness research: Was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions? For quality improvement: Was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle? For program evaluation: Was there a difference in obesity among people with diabetes who participated in the BMI reduction program? In addition, we provide an example from an published article to further explicate how to abstract information from the literature into the PIO MM literature matrix.

Abstract

Background and Purpose:

It is critical to accurately represent strengths interventions to improve data and enable intervention effectiveness research from a strengths perspective. However, it is challenging to understand strengths interventions from the multiple perspectives of computerized knowledge representation, evidence-based practice guidelines, and practice-based evidence narratives. Intervention phrases abstracted from nurse care coordinator practice narratives described strengths interventions with community-dwelling elders. This project aims were to (a) compare nurse care coordinator use of evidence-based interventions as described in the two guidelines (what to do and how to do it), (b) analyze nurse care coordinator intervention tailoring (individualized care), and (c) evaluate the usefulness of the Omaha System for comparison of narrative phrases to evidence-based guidelines.

Methods:

Phrases from expert nurse care coordinators were mapped to the Omaha System for comparison with the guidelines interventions and were analyzed using descriptive statistics. Venn diagrams were used to visually depict intervention overlap between the guidelines and the phrases.

Results:

Empirical evaluation of 66 intervention phrases mapped to 14 problems using 3 category terms and 19 target terms showed alignment between guidelines and the prhases, with the most overlap across two guidelines and phrases in Categories, and the most diversity in care descriptions.

Conclusion:

These findings demonstrate the value in having both standardized guidelines and expert clinicians who see the whole person and can synthesize and apply guidelines in tailored ways. There is potential to create a feedback loop between practice-based evidence and evidence-based practice by expanding this approach to use of practice-generated Omaha System data as practice-based evidence. Further research is needed to refine and advance the use of these methods with additional practices and guidelines [2].

In this example, we assume that this article is of interest for your project. It is relevant to PIO MM because it describes interventions, which are one of the PIO MM concepts. You will have identified several articles or other sources based on their titles and abstracts. Each source may have relevance for one or more concepts in the PIO MM.

To begin completing the PIO MM literature matrix for your project (Worksheet A), select one source, review it to make sure that it is relevant, and if so, read the entire article with the PIO MM concepts in mind. Then begin adding data for that source into one row in the PIO MM matrix. This one row that will be used only for that source.

10.2.1 Population of Interest

Descriptions of the population of interest may be found in reports from governmental or disease/condition-specific organizations, as well as the scientific literature. For example, for intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions*? For quality improvement: *was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* For the strengths-care coordination study example: the abstract describes the population of interest as *community-dwelling elders*. The population of interest was further explicated as follows: "In this project, we evaluated practice-based evidence from the newly emerging strengths approach used to promote well-being among adults with chronic conditions [2, p. 41].

Searches should focus on the populations of interest in each example. Information about high risk mothers may be found under a number of key words: e.g. mother, woman, women, infant, child, risk factors, home visiting, MCH, maternal-child health; information about people who smoke may be found under key words: e.g.

smoker, cigarette, tobacco, tobacco cessation, tobacco use; and searches for people with diabetes who may be obese may be found under key words: e.g. diabetes, diabetes mellitus, diabetes type 2, obesity, metabolic syndrome. Identify several sources that describe the population of interest from various perspectives and note this demographic information in the matrix. Use one row per source to abstract the data, and fill all applicable columns.

10.2.2 Problem Addressed

Descriptions of the extent of the problem may be found in reports from governmental or disease/condition-specific organizations, as well as the scientific literature. For example, for intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions*? For quality improvement: *was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* For the strengths-care coordination study example: the problem addressed is the alignment of practice-based interventions with evidence-based practice interventions, as in the title *Care coordination from a strengths perspective: A practice-based evidence evaluation of evidence-based practice* [2].

While similar key words as described for the population of interest may be used, the goal of this search is to document the extent of the problem in terms of cost to society and human suffering. What are the health disparities experienced by high risk mothers and what are the impacts of these disparities? For intervention effectiveness research, it is critical to describe health disparities of the particular population of interest in terms of previous research. What is the extent of smoking and what are the consequences of continuing to use tobacco? In particular for quality improvement it is critical to emphasize the impact of increased utilization of health care due to tobacco use on the health system, as well as the broad impact of tobacco use on morbidity and mortality. What is the population health impact of obesity among people with diabetes? To justify the importance of a program that addresses obesity among people with diabetes, it is critical to show how people with diabetes who are obese have poorer health outcomes than those who are not obese. Identify several sources that describe the problem from various perspectives and note this demographic information in the matrix. Use one row per source to abstract the data, and fill all applicable columns.

Note that strengths-care coordination study example is an evaluation of interventions as described by narratives and standards—not an intervention-outcome analysis, and thus there is no measure of outcome as described in the next section [2].

10.2.3 Measure(s) of Outcome

Literature or reports that describe the problem and intervention may also describe outcomes of the intervention. Measures of the Problem with psychometric

properties may be found in descriptions of instrument development. Noting the way in which outcomes are reported (e.g. benchmark attainment, improvement) will assist in planning how measures may be used in your project. For example, for intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions*? For quality improvement: *was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* To search for outcomes measures in intervention effectiveness research, quality improvement activities, and program evaluation, you may find any of the previously mentioned key-words useful; and in addition, add the key words outcome, measure, and instrument. Identify several sources that describe the outcome measures, including change scores (P_{Time2}-P_{Time1}) and benchmarking (P_{Time2}), as well as analysis methods, and note in the matrix. Use one row per source to abstract the data, and fill all applicable columns.

10.2.4 Intervention(s) Used

Descriptions of the intervention may be found in primary reports, published intervention manuals, evidence-based guidelines, or other local sources, such as agency procedures. Primary reports or systematic reviews of other effectiveness studies or efficacy studies for the same procedure, intervention, or program that provide rationale for addressing the problem using the intervention may be found in the scientific literature or other credible sources such as governmental guidelines websites or disease/condition-specific organizations. For example, for intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions*? Public health nurse home visiting guidelines have been described in the literature [5]. For quality improvement: *was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* The description of the intervention may originate as evidence-based practice from a guideline or the scientific literature [6]. For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* A program description or training manual may be referenced [7]. For the strengths-care coordination study example abstract: Intervention phrases abstracted from nurse care coordinator practice narratives described strengths interventions with community-dwelling elders [2]. This is further explicated in the article:

> The advent of computerized documentation offers opportunities to formally structure and document interventions. One such method is the use of a standardized terminology that uses defined terms to convey health care information at the user interface. From an information technology and data management perspective, the use of terminologies is essential for rigorous knowledge representation. Previous studies have shown that a strengths-based perspective to interventions can be documented using standardized terminologies. The use of a standardized terminology such as the Omaha System to represent clinical guidelines in EHRs allows for seamless translation of evidence into practice. Two clinical guidelines that are specific to com-

> munity-dwelling primary care patients/older adults with chronic conditions have been developed using the Omaha System: Community-Dwelling Elders (CDE) and Strengths Interventions [2, pp 41-42] …The project reused 66 phrases from existing de-identified descriptions of strengths based nursing practice (from) nurse care coordinator narratives [2, p 44].

For intervention effectiveness research, use key words as described for the population of interest and problem, adding key words related to the interventions and interventionist (e.g. public health nurse, intervention, fidelity, measure). For quality improvement, add key words regarding the proposed intervention change (e.g. evidence-based practice, health system, quality, QI). For program evaluation, include key words regarding the program and setting, and include words used in community settings and interventions (e.g. program, overweight, weight, blood sugar, school, community). Identify several sources that describe the intervention from various perspectives and note this intervention information in the matrix. Use one row per source to abstract the data, and fill all applicable columns.

10.2.5 Measures of Intervention

Measures of intervention are related to the intervention description above. For example, for intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions*? Public health nursing interventions were documented in the EHR during routine practice using the Omaha System [4, 5]. For quality improvement: *was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* The QI PDSA cycle is a period of time during which the evidence-based intervention was implemented, and the number of times that a person who smoked cigarettes on admission to care received the intervention during the QI PDSA cycle may be counted [6]. For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* The intervention may be measured by the proportion of the program completed by each person with diabetes and obesity. For the strengths-care coordination study, the measure of intervention was the Omaha System: This is further explicated in the article: The phrases were categorized according to the Omaha System Problem Classification Scheme (Problem), and Intervention Scheme (Category, Target, and care description) [2, p. 44]. The psychometrics of the Omaha System were not reported in the article, but this detail, if provided, would be appropriate in this section of Worksheet C. Use one row per source to abstract the data, and fill all applicable columns.

10.2.6 Measure of Intervention Fidelity

Measures of intervention fidelity may be difficult to identify in publications, so it is helpful to think broadly. For example, for intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions*? Public health nursing

interventions may be compared to expected interventions relative to the guidelines [3–5]. For quality improvement: *was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* Fidelity to the QI PDSA cycle may be measured using EHR data by interventionist and encounter [6]. For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* Intervention fidelity may be measured by the extent to which the program was delivered according to specified program elements and presenter credentials [7]. For the strengths-care coordination study, no specific information about care coordination program or interventions was provided, however, the article reported that "the nurse care coordinator narratives were provided by two coauthors who were bachelor's- or master's-prepared registered nurses with strengths-based training and an average of 10 years of experience in nurse care coordination" [2 p. 46].

10.2.7 Demographic Characteristics of a Sample

The characteristics of the sample are important for comparison across studies. For example, for intervention effectiveness research: was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions? Describe the number of high risk mothers and measures of central tendency such as age, race, ethnicity, marital status [5]. For quality improvement: was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle? Describe the number of people who smoked cigarettes on admission who received the QI intervention, and measures of central tendency such as age, race, ethnicity [6]. For program evaluation: Was there a difference in obesity among people with diabetes who participated in the BMI reduction program? Describe participants in the BMI reduction program including number of participants and measures of central tendency for BMI, age, sex, race, marital status, and ethnicity [7]. For the strengths-care coordination study, the sample consisted of strengths phrases in narratives of care coordinators who were bachelor's- or master's-prepared registered nurses with strengths-based training and an average of 10 years of experience in nurse care coordination [2 p. 44].

10.2.8 Contextual Factors

Recording information from each report or article will enable the evaluation of contextual factors across studies, especially in relationship to the context of your project. It is important to keep in mind that in intervention effectiveness research, quality improvement activities, and program evaluation the real world context of the problem is a fundamental and critical aspect of each type of evaluation. For example, For intervention effectiveness research: *Was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions*? Search for the context of public health nursing services (e.g. public health, department, agency) and qualifications (e.g. credential, degree); as well as

context of intervention delivery (e.g. home, community, visit, home visiting, family home visiting). For quality improvement: *Was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* Search for the context of the intervention (unit, hospital, clinic, outpatient) and the individuals who are involved (e.g. physician, provider, nurse practitioner, nurse, physician assistant, credential, degree). For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* For intervention effectiveness research) search for the previous key words regarding the program and context, and also search for interventionist key words (e.g. community health worker, community educator, health educator, public health nurse, degree, credential). For the strengths-care coordination study, the article notes the importance of the "real-life care coordination context, there is a critical need to evaluate evidence-based practice care guidelines based on real-world experiences of expert clinicians and patients" [2, p. 41]; but does not detail characteristics of the context. Identify several sources that describe contextual factors and note in the matrix. Use one row per source to abstract the data, and fill all applicable columns.

10.2.8.1 Contextual Factors – Interventionist

The characteristics of the interventionists are contextual factors that are important for comparison across studies and disciplines. For example, for intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions*? Document the credentials of public health nurses providing interventions [5]. For quality improvement: *was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* Document the credentials of clinicians providing the QI intervention [6]. For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* Document the program interventionist credentials [7]. For the strengths-care coordination study, the characteristics were limited to care coordinator credentials and experience in care coordination. Use one row per source to abstract the data, and fill all applicable columns.

10.2.9 Analysis Methods

Recording analysis methods of studies will provide clues for potential approaches that may be useful in your project. It is important to keep in mind that in intervention effectiveness research, quality improvement activities, and program evaluation methods options are highly varied and may include a number of different strategies in a single project as described in Chapters 5, 6 and 7. For example, in the case of the strengths-care coordination study, the methods approach consisted of using "descriptive statistics relative to [intervention] fit within and across the guidelines. Venn diagrams were used to visually depict the size of the overlap between guidelines and phrases. The analysis was conducted relative to Problem, Category, and Target terms across all sources. [2, p. 46]" This example demonstrates typical statistical approach (descriptive

statistics) used with a less common visualization technique (Venn diagrams). Use one row per source to abstract the data, and fill all applicable columns. Due to the diversity of approaches, it may be necessary to add columns to the matrix as needed to improve the synthesis of information regarding specific methods of interest in the study. Use one row per source to abstract the data, and fill all applicable columns.

10.2.10 Comments

The comments field is provided for recording additional information that may be useful. In the comments field, add any information that seems unusual or relevant to an aspect of your project that does not fit into another column. If you see trends in the comments field indicating information that should be added from all sources, add a column to the matrix. This will aid in synthesizing content across all sources.

10.2.11 Complete Reference

The complete reference for each source should be entered in the same row with all other data in its own column, using the style that will be used in the reference list of your project work. For example, the complete reference for the strengths-care coordination article (in APA style) [8] is presented as follows:

> Monsen, K. A., Vanderboom, C. E., Olson, K. S., Larson, M. E & Holland, D. E. (2017). Care Coordination From a Strengths Perspective: A Practice-Based Evidence Evaluation of Evidence-Based Practice. *Research and Theory for Nursing Practice, 31*(1), 39–55.

10.3 Sources of Information for the PIO MM Matrix

Literature for the PIO MM matrix may be identified in a number of sources. In particular, the Centers for Disease Control and Prevention [9], National Academies of Sciences, Engineering and Medicine-Health and Medicine Division Publications [10] and World Health Organization Fact Sheets and Reports [11] will likely yield information about population statistics and high level perspectives on health issues, economic impact, and intervention approaches. The United States Department of Health and Human Services Agency for Healthcare Research and Quality's National Guideline Clearinghouse is a public resource for summaries of evidence-based clinical practice guidelines [12]. These reports may be accessed free of charge through the organizational web sites.

Formal scientific literature searches may be conducted within subscription databases of the scientific literature, such as Scopus, Medline/Pubmed, CINAHL, PyscINFO, and the Cochrane Library [13–18]. Students, researchers, educators, and practitioners may access these databases through institutional accounts such as university library systems and governmental health department libraries (Table 10.1).

Scientific literature may also be identified using the internet through freely available browsers such as Google Scholar [19]; however there may be a charge for

Table 10.1 Sources for information needed to complete the PIO MM literature matrix

Database	Focus	Type of literature	Maintained by
Health and Medicine Division (HMD) formerly Institute of Medicine (IOM)	Science, technology, and medicine	Publications from the National Academies of Sciences, Engineering, and Medicine provide objective and straightforward advice to decision makers and the public	National Academies Press (NAP) was created by the National Academy of Sciences to publish the reports of the National Academies of Sciences, Engineering and Medicine, operating under a charter granted by the Congress of the United States
World Health Organization (WHO)	International public health issues: e.g. communicable disease, nutrition, occupational health, substance abuse	WHO reports available through the WHO global digital library	United Nations Development Group
Scopus (55 Million records)	Natural sciences, social sciences, arts, and humanities	Peer-reviewed literature: scientific journals, books and conference proceedings	Scopus Content Selection and Advisory Board to Elsevier
Medline/Pubmed (26 million records)	Biomedicine and health, covering portions of the life sciences, behavioral sciences, chemical sciences, and bioengineering	Biomedical literature from MEDLINE, life science journals, and online books	National Center for Biotechnology Information (NCBI), at the U.S. National Library of Medicine (NLM), located at the National Institutes of Health (NIH)
PsychINFO (4 million records)	Psychology and the behavioral and social sciences	Behavioral and social science research, dissertations, and scholarly literature abstracts	EBSCO Information Services platform

Table 10.1 (continued)

Database	Focus	Type of literature	Maintained by
CINAHL (3.6 million records)	Nursing, biomedicine, health sciences librarianship, allied health, alternative/ complementary medicine, consumer health	Journals, legal cases, clinical innovations, health care books, nursing dissertations, selected conference proceedings, standards of practice, and book chapters	CINAHL Information Systems, a division of EBSCO Information Services
Cochrane (10,000 reviews and protocols)	Evidence to inform healthcare decision-making dedicated to making up-to-date, accurate information about the effects of healthcare readily available worldwide	A database of systematic reviews and meta-analyses which summarize and interpret the results of medical research	Cochrane is an international not-for-profit and independent organization, The Cochrane Library interface is provided by Wiley InterScience

reading a full article if copyright is owned by the publisher. Google Scholar enables searches across disciplines from academic publishers, professional societies, online repositories, universities and other web sites. Scientific articles, theses, books, abstracts and court opinions and other relevant work not indexed in a subscription database may be identified using this method [19].

Completing the theory-based literature matrix provides a solid foundation regarding what is known in related to the population, problem, intervention, and context related to your project. Note that all rows may contain important information for more than one of the PIO MM concepts. Saturation of the matrix may be judged sufficient based on having substantive content in the columns that provide the foundation for your project. The following chapters build on this literature matrix in a step-by-step manner that will facilitate planning, implementation, analysis, and dissemination of the proposed PIO MM project.

Reflection Questions

- How does theory-based literature review enable operationalization of the PIO MM? Explain use of a theory together with the PIO MM to support a proposed project design and analysis.
- How does problem and definition of problem relate to outcome?
- What contextual factors are most important for your project? Consider setting, interventionist, and any unique factors related to your project.
- How does a literature matrix enable synthesis of the literature?

References

1. Gerrard J (2010) Health sciences literature review made easy: the matrix method, 3rd edn. Jones & Bartlett Learning, Burlington, MA
2. Monsen KA, Vanderboom CE, Olson KS, Larson ME, Holland DE (2017) Care coordination from a strengths perspective: a practice-based evidence evaluation of evidence-based practice. Res Theory Nurs Pract 31(1):39–55
3. Omaha System Guidelines. Intervention library: family home visiting. [Internet]. [cited 12 May 2017]. Available from http://omahasystemguidelines.org/?p=303
4. Martin KS (2005) The Omaha System: a key to practice, documentation, and information management, Reprinted 2nd edn. Health Connections Press, Omaha, NE
5. Monsen KA, Swenson SM, Klotzbach LV, Mathiason MA, Johnson KE (2017) Empirical evaluation of the changes in public health nursing interventions after the implementation of an evidence-based family home visiting guideline. Kontakt 19(3):e75–e85
6. Institute for Healthcare Improvement. The science of improvement: how to improve. [Internet]. [cited 12 May 2017]. Available from: http://www.ihi.org/resources/pages/howtoimprove/scienceofimprovementhowtoimprove.aspx
7. Centers for Disease Control and Prevention [US] (1999) Framework for program evaluation in public health. MMWR 48(RR11):1–40
8. American Psychological Association (APA) (2009) Publication manual of the American Psychological Association, 6th edn. American Psychological Association, Washington, DC
9. Centers for Disease Control and Prevention. Data and statistics. [Internet]. [cited 12 May 2017]. Available from https://www.cdc.gov/datastatistics/
10. National Academies of Sciences, Engineering, and Medicine. Health and Medicine Division Publications. [Internet]. [cited 12 May 2017]. Available from http://www.nationalacademies.org/hmd/Reports.aspx
11. World Health Organization (WHO) [Internet]. [cited 12 May 2017]. Available from http://www.who.int/en/
12. National Guideline Clearinghouse [US DHHS] [Internet]. [cited 12 May 2017]. Available from https://www.guideline.gov/
13. Scopus [Internet]. [cited 12 May 2017]. Available from https://www.elsevier.com/solutions/scopus
14. Pubmed [Internet]. [cited 12 May 2017]. Available from https://www.ncbi.nlm.nih.gov/pubmed?otool=igbcambulib
15. CINAHL [Internet]. [cited 12 May 2017]. Available from https://www.ebscohost.com/nursing/products/cinahl-databases/cinahl-complete
16. PsychINFO [Internet]. [cited 12 May 2017]. Available from http://www.apa.org/pubs/databases/psycinfo/
17. Medline [Internet]. [cited 12 May 2017]. Available from https://www.nlm.nih.gov/bsd/pmresources.html
18. Cochrane Library [Internet]. [cited 12 May 2017]. Available from http://www.cochranelibrary.com/
19. Google Scholar [Internet]. [cited 12 May 2017]. Available from http://scholar.google.com/

Define the Problem (Worksheet B)

11

11.1 Preparing to Complete Worksheet B

In intervention effectiveness research, quality improvement activities, and program evaluation, *the problem is the issue, topic, or risk factor that the intervention addresses*. This is the first and most essential step in intervention effectiveness research, quality improvement activities, and program evaluation.

Often this step is overlooked because the question on everyone's mind is "Did we make a difference?" However, before that question can be answered, it is critical to clearly define the missing and implied problem that the intervention addresses: "Did we make a difference in _____?" For intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions* (the problem is lack of health equity, or having health disparities)? For quality improvement: *was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle* (The problem is smoking cigarettes)? For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program* (the problem is obesity)? The fictitious study used to illustrate worksheet completion in this chapter examines variable health outcomes of older adults associated with frailty [1] and social determinants of health [2, 3]. The references in the worksheets [1–13] are intended as examples of the types of references that may be relevant to the problem and methods described in the worksheets for the fictitious study.

Interventions may address many problems at once, as in public health nurse home visiting programs that provide complex social interventions to promote optimal life-course development among high risk, first-time, single mothers. The 'Did we make a difference?' question may refer to a number of health and social problems of this population, such as poverty, domestic violence, depression, substance abuse, and more [14–18]. The complexity of understanding intervention effectiveness and outcome for

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_11

such situations underscores the need for clear definition of the problem(s) and associated measures. This is true in home care as well, with the notion of frailty added to the complex chronic conditions faced by older adults [1].

Defining the problem will enable documentation of measures that describe the problem and analysis of change in the problem over time, before and after intervention. Worksheet B may be used to define the problem. Review the literature matrix for the information that is needed to define the problem. One or more studies may provide the essential information, and if gaps remain in the problem definition, further literature review may be necessary. In order to clearly link the problem as described in Worksheet B and measure(s) that operationalize the problem, Worksheet B (problem definition) will inform the completion of Worksheet D (outcome measure for the problem).

Directions to assist in completing Worksheet B are provided here. Additional details about defining the problem are available in Chapter 2. Use the completed PIO MM literature matrix as a basis for completing the worksheet. Keep in mind that you are completing this worksheet and all subsequent worksheets for your project, and include information from the matrix that is relevant. For example, several problems may be described in the PIO MM matrix. Choose literature for Worksheet B that most closely and rigorously defines and describes the problem that will be addressed in your project.

For example, related to the question for intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?* The gap in knowledge may be stated as: Little is known about differences in health equity after public health nursing interventions. For quality improvement: *was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* The gap in knowledge may be stated as: It is not known to what extent people in who smoked cigarettes on admission to care changed cigarette smoking behavior. For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* The gap in knowledge may be stated as: It is not known to what extent obesity decreased among people with diabetes who participated in the program. Add to this statement the way that the outcome of interest will be calculated.

11.2 Step-by Step Instructions for Completing Worksheet B

11.2.1 Problem

At the top of Worksheet B, state the problem that will be defined in the worksheet. For instance, the sample worksheet included later in this chapter includes this problem statement: “Variable health outcomes of older adults associated with Frailty and Social Determinants.”

11.2.2 Definition of the Problem

In the first row of the worksheet, define the problem and provide a reference for the definition. For example, For intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?* Give the definition of health equity and reference for that definition. For quality improvement: *was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* Give the definition of cigarette smoking and reference for that definition. For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* Give the definition of obesity and reference for that definition. For the fictitious study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions, the problem is poor health outcomes that may be improved by home care nursing. In this case, the worksheet includes the definition of the problem as follows:

Poor outcomes of older adults are a burden to the health care system and may be related to frailty, defined as physiological decline in late life, characterized by marked vulnerability to adverse health outcomes.[1] and may also be associated with the social and behavioral determinants of health as defined by the Institute of Medicine [2, 3].

11.2.3 Population of Interest

Next, describe the population of interest: the particular community, group, subgroup, or type of people that experience the problem or are at risk of experiencing the problem addressed by the intervention. Use the literature review to populate this section through synthesis of the population characteristics described in the PIO MM matrix. For the fictitious study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions, the population of interest is older adults receiving home care interventions. There is potential to further define this group by age or other demographic characteristics that may be of particular interest.

11.2.4 Background

Next, complete the background information regarding the problem and population. Several prompts are provided to assist in substantiating the importance of the problem based on the literature, including statistics about the percentage or number of individuals in the total population with the problem or are at risk, costs incurred addressing the problem, years of potential life lost due to the problem, and health system-related gaps in addressing the problem.

For example, for intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?* Describe health disparities for high risk mothers with references. If little is known, describe health disparities for a related population, such as women in poverty. For quality improvement: *was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* Describe the negative impact of smoking on health, and the trends in numbers of smokers among the population of interest, with references. Use literature that most aligns with the population served by the health system or local clinic. For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* Describe the detrimental health effects of diabetes, in particular related to obesity, and the trends in numbers of individuals with diabetes, pre-diabetes, and obesity in the population, with references. As suggested in Chapter 10, to substantiate the compelling need to address the problem, use highly respected sources and reviews of the literature to substantiate the problem.

These prompts in Worksheet B are suggestions and are not intended to be a complete list. Other background information may be essential to your project; and more columns to describe relevant background (contextual information) be added to Worksheet B as needed. For the fictitious study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions, several brief suggestions are provided related to the clinical and policy aspects of home care interventions and outcomes for the population of interest.

11.2.5 Problem Measurement Instrument/Scale

Next, describe the Problem measurement instrument/scale (s). In Worksheet B, note the measure(s), instrument(s), and scale(s) that operationalize the problem concepts. In Chapter 13, you will be guided to compete a worksheet for each measure. Including the measures in Worksheet B enables consistency in planning across the worksheets. For intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions*? Describe specifically how health equity will be measured, with references. *For quality improvement: was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle*? Describe how smoking behavior will be measured, with reference. For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* For example, *did we make a difference in the behavior of people who smoked cigarettes on admission to care?* Describe specifically how obesity will be measured, with references. For the fictitious study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions, the existing dataset that measures health care outcomes was generated by practicing home care clinicians (nurses and physical therapists) during the routine documentation of care using the Omaha System Problem Rating Scale for Outcomes [9].

11.2.6 Anticipated Outcome and Rationale

Given that the study will examine an efficacious intervention or a proven program, complete the anticipated outcome and rationale for the study. For example, for intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?* Describe the expected change in health equity with reference. *For quality improvement: was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* Describe the expected change in smoking with references. For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* Describe the expected change in obesity, with references. For the fictitious study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions, the use of evidence-based home care interventions has been described in the literature and online [10, 11].

11.2.7 What is Not Known/Gap in Knowledge

Finally, summarize what is not known—the gap in knowledge as defined previously: It is not known if the intervention addressing the problem that is (efficacious or effective) for a given population in (a controlled environment or with other populations in real world settings) also will be associated with positive outcomes in the identified setting and population. This is because there are not studies in the literature specific to the gap we have defined. For example, for intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?* The gap in knowledge with measure may be stated as It is not known to what extent high risk mothers had changes in health equity. This change in health equity may be measured by the difference in health equity measure X on discharge compared to health equity measure X on admission. For quality improvement: *was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* The gap in knowledge with measure may be stated as It is not known to what extent people in who smoked cigarettes on admission to care changed cigarette smoking behavior. This change in behavior may be measured by the difference in the number of cigarettes per day now compared to number of cigarettes per day on admission. For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program*? The gap in knowledge with measure may be stated as It is not known to what extent obesity was decreased among people with diabetes This change in obesity may be measured by the difference in BMI now compared to BMI on admission. For the fictitious study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions, the gap in knowledge was described as follows:

> Little is known about variability associated with frailty and social determinants in the health outcomes of older adults receiving home care services.

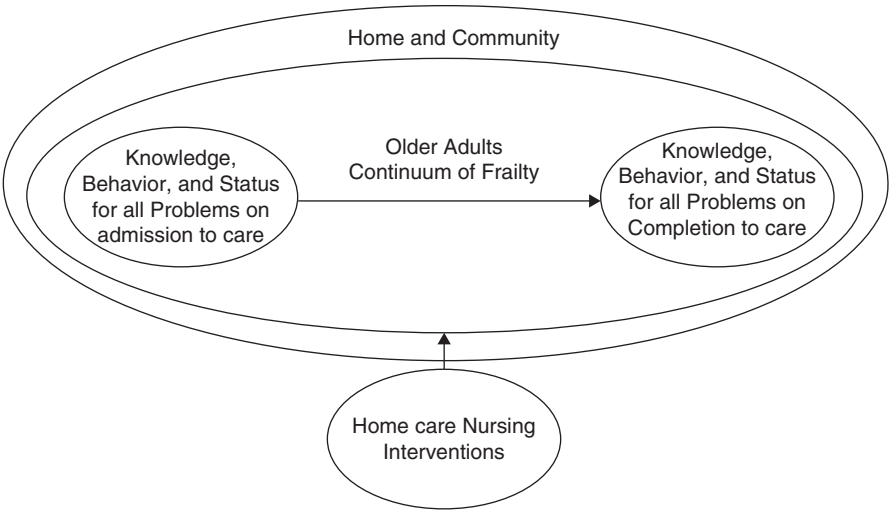

Fig. 11.1 PIO MM for evaluation of health outcomes among older adults receiving home care interventions: frailty and social determinants. This example is loosely based on a study by Grace Gao, Sasank Maganti, Kari Miller, and Karen A. Monsen (in progress). The results are fictional and are intended to illustrate the types of data analyses that explore and demonstrate care quality and outcomes for groups within an existing dataset.

Upon completion of Worksheet B, the reader will have defined the problem, substantiated the importance of the problem with relevant literature, and identified the method for measuring the problem over time. After completing Worksheet B it may be helpful to share the worksheet with colleagues and clinical experts for review and comment, to ensure that the worksheet content correctly addresses the issue as identified in the real-world setting, aligns with the literature, and is internally consistent across all content. Completing the PIO MM diagram for the problem and other PIO MM concepts is recommended at the completion of Worksheet B, and may be revised throughout the study planning process (Fig. 11.1).

Reflection Questions

- What is the value and/or importance of defining a problem according to the literature?
- How does the population of interest relate to the problem?
- In the ideal intervention effectiveness research, quality improvement activities, or program evaluation project, what design and measures should be used to evaluate the outcomes? Is this possible? If not, what alternatives are available?
- How is the identified gap in knowledge substantiated in the literature?

Worksheet A. References from the Preliminary Literature Review

- Fried, L.P., Tangen, C.M., Walston, J., Newman, A.B., Hirsch, C., Gottdiener, J., Seeman, T., Tracy, R., Kop, W.J., Burke, G., McBurnie, M.A. (2001). Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci., 56(3): M146-56.[1]
  - This study provides a basis for mapping existing data to a measure for evaluating the associations between frailty and home care outcomes
- Institute of Medicine. (2014a). Capturing social and behavioral domains in electronic health records: Phase 1. Washington, DC: The National Academies Press.[2]
- Institute of Medicine. (2014b). Capturing social and behavioral domains and measures in electronic health records: Phase 2. Washington, DC: The National Academies Press.[3]
  - This series of articles provides the basis for mapping existing data to recommended social determinants measures
- Monsen, K.A., Brandt, J.K., Brueshoff, B.L., Chi, C., Mathiason, M.A., Swenson, S.M., & Thorson, D.R. (2017). Social Determinants and Health Disparities Associated With Outcomes of Women of Childbearing Age Who Receive Public Health Nurse Home Visiting Services. JOGNN, 46, 292–303. [4]
  - This study provides a basis for the use of existing nurse-documented assessments in the study of social and behavioral determinants of health relative to overall outcomes of a community nursing program

Additional literature is reflected in the references within the other worksheets.

Worksheet B. Definition of the Problem with Statement of the Gap in Knowledge

Problem: Variable health outcomes of older adults associated with frailty and social determinants

Definition of the problem	Poor outcomes of older adults are a burden to the health care system and may be related to frailty, defined as physiological decline in late life, characterized by marked vulnerability to adverse health outcomes.[1] and may also be associated with the social and behavioral determinants of health as defined by the Institute of Medicine [2, 3]
Population of interest	Older adults receiving home care services who may be frail as defined by Fried et al.[1] and have outcomes associated with social and behavioral determinants of health as defined by the Institute of Medicine [2, 3]
Background: Literature describing the problem (% or number of individuals in the total population)	Growing population of older adults needing home care services [6, 7]
Background: Costs incurred addressing the problem)	Total cost of home care services $62 Billion public expenditures (Medicaid and Medicare) in 2008 [7]
Background: Years of potential life lost due to the problem	

Background: Health system-related gaps in addressing the problem	Policy changes affecting availability of home care [8]
Problem measurement instrument/scale	Variable health outcomes of older adults receiving home care services as measured by Omaha System Problem Rating Scale for Outcomes on admission and dismissal from Home Care services [9]
Anticipated outcome and rationale	It is critical to understand the needs for services and home care intervention effectiveness for the most vulnerable older adults to provide data for policy makers and health system leaders. It is likely that home care interventions improve health outcomes of older adults [6, 7]. There may be patterns in the health outcomes of older adults receiving home care services. Previous studies have shown variability in health outcomes of other populations that were associated with social and behavioral determinants [4] . For older adults, frailty is also an important factor [1] that should be examined
What is not known/gap in knowledge	Little is known about variability associated with frailty and social determinants in the health outcomes of older adults receiving home care services

Worksheet C. Describe the Intervention Used to Address the Problem

Definition of the intervention	Home care interventions known to have occurred between admission and discharge—Evidence-based home care guidelines [10–13]
Percentage of persons who improved after intervention	Home care patients improved after intervention [6, 7, 12]
Meta-analysis showing levels of effectiveness across studies	There is evidence of effectiveness of home-based health promotion interventions for older adults [13]
For whom	Differences between frail and non-frail elders [12] Studies of patients with chronic or incurable disease, fractures, neurological conditions and cardiorespiratory conditions [13]
Under what conditions	Home care [10–13]
Theory of causal mechanism	None reported
Essential (core) components	Evidence-based home visiting interventions [10, 11] home-based, nurse-led health promotion [13]
Intervention content	Tailored to meet individual needs [12]
Intervention adherence and involvement	Not available in the dataset
Intervention measure—amount	Not available in the dataset
Intervention measure—type	Not available in the dataset
Intervention measure—fidelity	Not available in the dataset
Intervention measure—quality	Not available in the dataset
Interventionist—qualifications	Not available in the dataset
Interventionist—training	Nurses and physical therapists [12, 13]
Interventionist—demographics	Not available in the dataset
Interventionist—organization	De-identified merged dataset from numerous organizations
Source of data	Existing home care dataset

Worksheet D. Define the Outcome and Related Measures or Scales

Definition of the problem	Problem-specific outcomes of older adults receiving home care services
Measure/scale that operationalize the problem	Omaha System Problem Rating Scale for Outcomes Knowledge, Behavior, and Status Scales [9]
Use of measure in previous research	Extensive use in previous intervention effectiveness research, quality improvement activities, and program evaluation, especially in public health settings [4, 9, 12]
Psychometric properties	Problem Rating Scale for Outcomes is a valid, reliable measure with psychometric properties established during development and subsequently during application in practice and research [9]
Validity (construct)	Described in Martin [9]
Reliability (internal consistency)	Described in Martin [9]
Reliability (test-retest)	Described in Martin [9]
Reliability (across raters)	Not available in the dataset
Data collection—strategy	Home care nurses and physical therapists assessed and documented the KBS ratings during routine care
Data collection—training	Not available in the dataset
Data collection—timing	Admission to care and dismissal from care
Source of data	Existing home care dataset

Worksheet E. Plan the Analysis Methods

Exploratory data analysis: Sample characteristics	Patterns in the sample related to problems by social determinants metric and by frailty on a continuum of values using heat maps
Exploratory data analysis: Interventions	No EDA planned with intervention data
Exploratory data analysis: Outcomes	Patterns in outcomes by social determinants and frailty using line graphs
Descriptive analysis: Sample	Means, SD, Frequencies
Descriptive analysis: Interventions	Length of intervention in days (see Time analysis)
Descriptive analysis: Problems	Frequencies of problems by group
Descriptive analysis: Outcome P_{Time1}, P_{Time2}	Descriptive analysis of KBS outcomes overall, and by social determinants and frailty using means, SD, frequencies
Inferential analysis: Outcome P_{Time1}, P_{Time2}	Significance of KBS outcomes overall (P_{Time2}-PT_{ime1}) and by group using parametric or nonparametric tests
Inferential analysis: Effect size	Effect size of overall outcomes using Cohen's *d*
Inferential analysis: Benchmark attainment	Significance of KBS change outcomes overall (P_{Time2}-PT_{ime1}) and by group using parametric or nonparametric tests
Inferential analysis: Survival	No survival analysis planned
Time analysis	Time is reflected by before and after time-points—mean of difference between dates, SD (days of service)
Inferential analysis: Correlations	Associations between length of services and baseline assessments of knowledge, behavior, and status
Qualitative analysis	Not available in the dataset

Worksheet F. Results Statements and Alignment with the Literature

Sample characteristics	The sample of 1613 older adults was evenly divided between individuals 65–80 (51%) and 81 or above, with a higher percentage of females (75%) overall and by age group (60% and 90%, respectively)
How sample aligns with literature or demographics of the population	Sample demographics represent the typical characteristics of individuals receiving home care [7]
Interventions	Length of service was positively associated with frailty ($r(1,13) = 0.35$, $p < 0.01$) and by social determinants scores ($r(1613) = 0.47$, $p < 0.01$)
How interventions align with the literature	This pattern has not been explored in the literature to date
Outcomes	A t-test was conducted to examine significance of improvement in K, B, and S after home care intervention. There was significant improvement overall for knowledge ($t(1613) = 4.43$, $p < 0.01$, $d = 0.51$), behavior ($t(1613) = 4.36$, $p < 0.01$, $d = 0.67$), and status ($t(1613) = 5.52$, $p < 0.01$, $d = 0.68$). A two-way ANOVA was conducted for each of three outcome variables (K, B, and S) on the influence of frailty and social determinants on overall outcomes of older adults receiving home care services. The main effect for social determinants groups on status scores yielded an F ratio of $F(1, 1613) = 127.2$, $p < 0.01$, indicating a significant difference between older adults with and without social determinants problems. The main effect for Frailty on status scores yielded an F ratio of $F(1, 1613) = 155.4$, $p < 0.01$. Indicating that the effect for social determinants was significant across scores, with decreasing final and change scores as frailty increased. Social determinants, Frailty, and interactions between Social determinants and Frailty were significantly associated with differences in Status outcomes ($F(2, 1136) = 13.41$, $p = 0.01$)
How outcomes align with the literature	This aligns with previous literature [12, 13] and adds to what is known about frailty and social determinants for older adults receiving home care. Findings are novel and should be replicated with other home care data
Relationship between interventions and outcomes	Length of service was negatively correlated with Knowledge scores on admission ($r(1316) = -0.36$, $p = 02$)
How intervention/outcome relationships align with the literature	This aligns with previous literature [12, 13] regarding the need for additional care when there are cognitive or health literacy issues

References

1. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Seeman T, Tracy R, Kop WJ, Burke G, McBurnie MA (2001) Frailty in older adults evidence for a phenotype. J Gerontol Ser A Biol Med Sci 56(3):M146–M157
2. Institute of Medicine [US] (2014) Capturing social and behavioral domains in electronic health records: Phase 1. National Academies Press, Washington, DC

3. Institute of Medicine [US] (2014) Capturing social and behavioral domains and measures in electronic health records: Phase 2. National Academies Press, Washington, DC
4. Monsen KA, Brandt JK, Brueshoff B, Chi CL, Mathiason MA, Swenson SM, Thorson DR (2017) Social determinants and health disparities associated with outcomes of women of childbearing age who receive public health nurse home visiting services. J Obstet Gynecol Neonatal Nurs 46(2):292–303
5. Medical Subject Headings (MeSH). Definition of health disparities. [Internet]. [cited 12 May 2017]. Available from https://www.nlm.nih.gov/hsrinfo/disparities.html
6. Institute of Medicine [US] (2015) Forum on aging. In: The future of home health care: workshop summary. National Academies Press, Washington, DC
7. National Association of Home Care and Hospice. Statistics 2009 [Internet]. [cited 12 May 2017]. Available from http://www.nahc.org/assets/1/7/10hc_stats.pdf
8. National Association of Home Care and Hospice. Key legislation affecting home care and hospice [Internet]. [cited 12 May 2017]. Available from http://www.nahc.org/advocacy-policy/key-legislation-affecting-home-care-and-hospice/
9. Martin KS (2005) The Omaha System: a key to practice, documentation, and information management. Reprinted 2nd edn. Health Connections Press, Omaha, NE
10. Omaha System Guidelines. Intervention library: family home visiting. [Internet]. [cited 12 May 2017]. Available from http://omahasystemguidelines.org/?p=303
11. Monsen KA, Foster DJ, Gomez T, Poulsen JK, Mast J, Westra BL, Fishman E (2011) Evidence-based standardized care plans for use internationally to improve home care practice and population health. Appl Clin Inform 2(3):373–383
12. Monsen KA, Westra BL, Oancea SC, Yu F, Kerr MJ (2011) Linking home care interventions and hospitalization outcomes for frail and non-frail elderly patients. Res Nurs Health 34(2):160–168
13. Tappenden P, Campbell F, Rawdin A, Wong R, Kalita N (2012) The clinical effectiveness and cost-effectiveness of home-based, nurse-led health promotion for older people: a systematic review. Health Technol Assess 16(20):1–72
14. Monsen KA, Chatterjee SB, Timm JE, Poulsen JK, McNaughton DB (2015) Factors explaining variability in health literacy outcomes of public health nursing clients. Public Health Nurs 32(2):94–100
15. Maternal infant and early childhood home visiting [US DHHS], 2016. [Internet]. [cited 12 May 2017]. Available from http://mchb.hrsa.gov/programs/homevisiting/index.html
16. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K (2011) Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 155(2):97–107
17. Nurse-Family Partnership. [Internet]. [cited 12 May 2017]. Available from https://www.nurse-familypartnership.org/ (2014)
18. Edwards N, Di Ruggiero E (2011) Exploring which context matters in the study of health inequities and their mitigation. Scand J Public Health 39(6 suppl):43–49

Describe the Intervention (Worksheet C) 12

12.1 Preparing to Complete Worksheet C

In intervention effectiveness research, quality improvement activities, and program evaluation, the interventions that are evaluated have been found to be efficacious in controlled settings or found to be effective with other populations in real-world settings. In this chapter the focus is on Worksheet C, describing the intervention. The example presented in this chapter is the evaluation of public health nurse home visiting interventions with improved outcomes of adult and adolescent mothers with and without the Mental health problem [1–15].

It is an expectation of leadership in healthcare organizations that data should be used to *link intervention processes to health outcomes* [16]. On the trajectory of research translation from bench to bedside, it is essential to test efficacious interventions in real life settings. The challenge is to deliver the intervention with fidelity to the original as designed and intended [17]. Thus, it is essential to describe the intervention or program and document fidelity to the intended intervention or program. Such interventions may be embedded within programs that are delivered as an 'evidence-based' intervention. For example, the Nurse-Family Partnership is a program that gave extensive attention to intervention fidelity as the intervention was translated into a national program after 20 years of research [18, 19].

Complex social interventions are multi-faceted, tailored to the needs of identified individuals within unique social situations [18–21]. Often these interventions are evaluated using a black-box approach [16, 20, 21]. A major concern of intervention effectiveness research, quality improvement activities, and program evaluation is that interventions may result in inconsistent effects despite being based on prescribed programs or robust theoretical frameworks [16, 20, 21]. It is even more challenging to evaluate interventions when best practices include intervention tailoring to meet unique individual needs [6, 16, 20, 21]. Thus, whenever possible it is advantageous to describe the interventions used to address the problem that was defined in Worksheet B.

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_12

Having intervention data enables the description of interventions that are differentially associated with change in outcomes. Understanding how and why such interventions bring about change and why they sometimes fail may be explored using the entire PIO MM with more advanced methods including multiple measures and mixed methods [7, 14, 16, 22].

From the perspective of program evaluation, process measures from the logic model (Inputs, Activities, and Outputs) describe resources, costs, actual tasks (interventions), and fidelity to planned intervention [23, 24]. From the perspective of nursing research, detailed intervention data from EHRs may be found in large nursing datasets to investigate questions of fidelity, tailoring, and theoretical alignment [4, 24–28]. From the perspective of quality improvement evaluation, detailed intervention information enables evaluation of care quality, relative to a situation-specific definition of care quality as operationalized using intervention data [5, 25–33].

12.2 Step-by Step Instructions for Completing Worksheet C

At the top of Worksheet C, state the intervention that will be described in the worksheet. In the first row, describe the intervention and provide a reference for the evidence-based intervention.

12.2.1 Describe the Intervention

For example for intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?* Describe public health nurse interventions in general, and for particular populations or problems, with references. *For quality improvement: was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* Describe of the smoking cessation intervention and a reference for that definition. For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* Describe the program that was used to address obesity, and reference for that program. For the fictitious study of intervention patterns associated with outcomes of adult and adolescent mothers with and without the Mental health problem, the intervention description is lengthy, incorporating the description of the Omaha System Intervention Scheme and its psychometric properties, examples of intervention terms, and the methods used to group interventions for the study.

12.2.2 Expected Effectiveness

The expected effectiveness of the intervention may be described in one or more of the following ways: the **percentage of persons who improved after intervention, meta-analysis showing levels of effectiveness across studies, for whom the**

intervention was effective, and under what conditions. This documentation provides the rationale for using the intervention in your proposed study. For example, describe the effectiveness or efficacy of the described intervention among other populations or settings, for example, a public health nursing intervention to improve health equity among older adults; a smoking cessation intervention effective with adults may be evaluated among older adults or adolescents; or an obesity screening intervention may be efficacious in the early childhood population in controlled research, but may not have been evaluated in a less-controlled community setting. This constitutes documentation of evidence-based practice. For the fictitious study of intervention patterns associated with outcomes of adult and adolescent mothers with and without the Mental health problem, the expected effectiveness information was provided from other family home visiting outcomes research.

12.2.3 Theoretical Framework

One or more **theoretical framework**(s) may be referenced as the basis of the intervention. Theory-based interventions have been found to be more effective than those not using theory [34–38]. The purpose of intervention effectiveness research, quality improvement activities, and program evaluation is to demonstrate outcomes, not test theoretical frameworks. However, if the intervention is theoretically based, it is essential to note theory as it may provide a **causal mechanism** that may aid in interpreting findings. This information should also be reported in any publications of the results in order to build the knowledge base related to the theoretical framework. A secondary study aim may be developed to address theory testing. If this is of interest, you may wish to consult the work of Acton and colleagues [36] as well as materials available on-line through the United States Office of Behavioral and Social Sciences Research [35]. For example, smoking cessation interventions may be based on the Transtheoretical Model of Change [37] while obesity screening interventions may be based on Pender's model [38]. For the fictitious study of intervention patterns associated with outcomes of adult and adolescent mothers with and without the Mental health problem, the study was based on the Clinical Nursing Models intervention effectiveness framework proposed by Barnard [8].

12.2.4 Intervention Content and Essential Core Components

Key information regarding intervention content and essential core components may be found in the theoretical framework description or in literature describing previous studies that used the intervention, and delineating the core content will support operationalizing the content in your project. Core components may include approaches to care such as motivational interviewing [39], or may be described in a more granular way as in EHR intervention data captured during documentation of care [5]. Related to public health nurse home visiting interventions, intervention content may be reported as counts of discrete interventions, or other measures of

intervention dose such as proportions of interventions by type [5, 7, 14]. Related to smoking cessation interventions, describe specifically how smoking cessation interventions measured and reported in your project, with references. For obesity among people with diabetes, describe specifically how obesity screening interventions will be measured and documented in your project, with references. For the fictitious study of intervention patterns associated with outcomes of adult and adolescent mothers with and without the Mental health problem, examples of core components of the evidence-based interventions were described in the literature and references were provided [5, 9, 10].

12.2.5 Describe Intervention Measurement: Amount, Type, Fidelity, Quality

The next several prompts are provided to assist in describing the intervention measures based on the literature, including how they may be used to report the intervention **amount, type, fidelity** and/or **quality**, as well as an individual's **adherence and involvement**. In particular, the standardized intervention terminologies described in Chapter 4 are among the most highly developed intervention measures, and there is a growing body of knowledge using these terminologies in intervention effectiveness research, quality improvement activities, and program evaluation [5, 26, 28–30]. Intervention amount may be reported using a yes/no (black box [16]), counts of defined components, length of time with interventionist or in a program. Type may be specific to the problem or theory or may be described by defined components as with EHR data. Fidelity is defined as the extent to which an intervention is deliver ed as intended [40]. Measures of intervention fidelity may be noted, with references [41]. For example, the use of an evidence-based care plan may be measured by evaluating the extent to which interventions in the care plan were documented in the EHR [10]. Quality of an intervention may depend on numerous factors such as available time and resources. If information is available regarding the quality of the intervention (as may be the case in quality improvement studies), the definition of a quality intervention should be noted with references. Given that the study will examine an efficacious intervention, a quality improvement project based on evidence, or a program with proven outcomes, provide information about the anticipated outcomes of the intervention and rationale for the study for the intervention. Information related to this for the identified problem may be found in Worksheet B, and should be consistent across worksheets. For example, related to smoking cessation interventions, describe the expected behavior change of people who smoked cigarettes based on previous studies, with references. For obesity screening interventions, describe the expected change in obesity, with references. For the fictitious study of intervention patterns associated with outcomes of adult and adolescent mothers with and without the Mental health problem, amount is the number of interventions; type is operationalized as Omaha System Intervention Scheme Category, and quality is supported by the use of evidence-based home visiting guidelines [10] as well as the rigor of the Intervention Scheme instrument [5, 12].

12.2.6 Describe Interventionist Characteristics: Qualifications, Training, Demographics

Next, consider characteristics related to the interventionist that may be important to your project, including **qualifications, training, demographics**, and **organization**. Such information may be needed to document (a) that the interventionist was equipped to deliver the intervention as described in the literature, or (b) your project may be able to stratify data or adjust for interventionist factors. For the fictitious study of intervention patterns associated with outcomes of adult and adolescent mothers with and without the Mental health problem this was derived from literature related to the study [4, 7, 9].

Finally note the source of data for your project, based on the measure(s) explained above. For example, there may be a publicly available dataset of interventions related to smoking cessation or obesity that is available for the population of interest [42, 43]. If your project will take place within a health care organization there may be existing EHR data that meets the requirements outlined in the literature.

Upon completing Worksheet C, the reader will have established a method of describing interventions within the proposed intervention effectiveness research, quality improvement activity, or program evaluation project. This method will guide the selection of the PIO MM design and analysis, whether from a black box [16] perspective or the granular description of each intervention activity [5, 26–30] and/or encounter. See the completed PIO MM diagram that shows relationships among the PIO MM concepts for this example (Fig. 12.1).

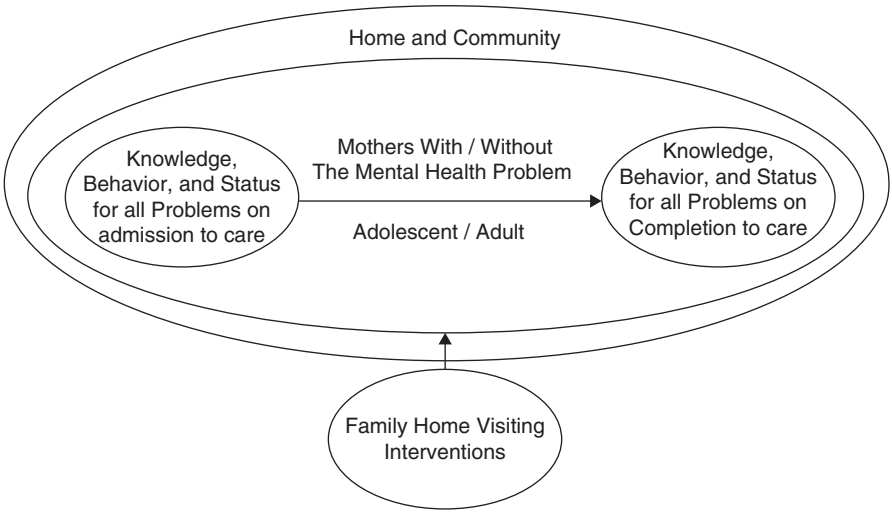

Fig. 12.1 PIO MM for evaluation of public health nurse home visiting interventions with improved outcomes adult and adolescent mothers with and without the Mental health problem. This example is loosely based on a study by Diane McNaughton, Amy B. Lytton, Young Shin Park, Carolyn Porta, Chih-Lin Chih, Michelle A. Mathiason, Lisa M. Moon, Joan K. Brandt, David M. Radosevich, and Karen A. Monsen. The results are fictional and are intended to illustrate the types of data analyses that explore and demonstrate care quality and outcomes for groups within an existing dataset

Reflection Questions

- In the ideal intervention effectiveness research or program evaluation project, what measures should be used to describe the intervention? Is this possible? If not, what alternatives are available?
- What is the suggested mechanism of action of the intervention according to the literature?
- What is the value and/or importance of measuring intervention fidelity?
- How can interventions be related to the problem measure in a statistical model?

Worksheet A References from the Preliminary Literature Review

- World Health Organization. Social determinants of mental health. 2014 http://apps.who.int/iris/bitstream/10665/112828/1/9789241506809_eng.pdf [1]
  - This World Health Organization publication describes the importance of mental health globally
- Goetzel RZ, Long SR, Ozminkowski RJ, Hawkins K, Wang S, Lynch W. Health, absence, disability, and presenteeism cost estimates of certain physical and mental health conditions affecting US employers. Journal of Occupational and Environmental Medicine. 2004 Apr 1; 46(4):398–412 [3].
  - This article gives an overview of the individual and societal burden of Mental health problems from the perspective of lost productivity and wage earnings
- Rahman A, Surkan PJ, Cayetano CE, Rwagatare P, Dickson KE. Grand challenges: integrating maternal mental health into maternal and child health programmes. PLoS Med. 2013 May 7; 10(5):e1001442 [2].
  - This article provides a review of the severe issues that mothers with Mental health problems face when seeking care
- Garcia, C., McNaughton, D., Radosevich, D.M., Brandt, J.K., Monsen, K.A. (2013). Family home visiting outcomes for Latina mothers with and without Mental health problems. Public Health Nursing, 30(5), 429–38 [4].
  - This study provides a description of the sample characteristics and outcomes that relate to the current study.

Additional literature is reflected in the references within the other worksheets.

Worksheet B Definition of the Problem with Statement of the Gap in Knowledge

Problem: Variability in risk related to mental health associated with poor health outcomes of mothers.

Definition of the problem	Mothers with Mental health problems are at risk for and suffer disproportionately from poor health and social outcomes [1–4]
Population of interest	Mothers with Mental health problems (particularly adolescent mothers)

Background literature describing the problem (% or number of individuals in the total population)	Mental health and many common mental disorders are shaped to a great extent by the social, economic, and physical environments in which people live [1] Social inequalities are associated with increased risk of many common mental disorders [1]
Background (costs incurred addressing the problem)	Costs extend beyond healthcare costs to societal costs in terms of lost productivity and earnings [1, 2]
Background (years of potential life lost due to the problem)	
Background (health system-related gaps in addressing the problem)	There are major gaps in access to mental health services for mothers with Mental health problems [3]
Problem measurement instrument/ scale	Health and social outcomes of the Omaha System using the Problem Classification Scheme and Problem Rating Scale for Outcomes [5, 14, 15]
Anticipated outcome and rationale	Improvement for all groups defined by adolescents and adults with and without the Mental health problem associated with PHN home visiting
What is not known?	What is differential intervention effectiveness of PHN home visiting for groups defined by adolescents and adults with and without the Mental health problem?

Worksheet C Describe the Intervention Used to Address the Problem

Describe the Intervention: Public Health Nurse Home Visiting Services.

Definition of the intervention	Preventive home visiting services that are tailored to promote optimal health and parenting outcomes [6, 7]. Interventions are defined according to the Omaha System [5]. Intervention Scheme describes four intervention actions (Categories) that may be used to address each of the 42 Problem concepts, together with 75 defined terms (Targets). Each intervention consists of one Problem, one Category, and one Target term, together with one customizable care description term [5]
Percentage of persons who improved after intervention	Women of childbearing age served by PHN home visiting programs show significant improvement [6, 7]
Meta-analysis showing levels of effectiveness across studies	
For whom	Adolescent and adult mothers with and without the Mental health problem [4]
Under what conditions	Receiving public health nurse family home visits
Theory of causal mechanism	Clinical nursing models middle range theory or theoretical framework derived from a randomized trial of maternal-child intervention effectiveness [8]
Essential (core) components	Interventions directed to address Omaha System problems common among high risk mothers with and without the Mental health problem such as Pregnancy, Postpartum, Family planning, Caretaking/parenting, Substance use, and Abuse [6, 7]
Intervention content	Standardized evidence-based interventions for family home visiting have been defined and incorporated within electronic health records used by public health nurses. Examples of interventions may be Pregnancy–Surveillance–signs/symptoms-physical–warning signs; and Postpartum–Teaching, Guidance, and Counseling–anatomy/ physiology–postpartum changes [5, 9, 10]

Intervention adherence and involvement	Not available in dataset
Intervention measure—Amount	Counts of Omaha System interventions per individual
Intervention measure—Type	Counts of Omaha System interventions by type based on the four Categories of the Omaha System: Teaching, Guidance, and Counseling, Treatments and Procedures, Case Management, and Surveillance [5]. Proportions of Categories for each individual used for round up—Round down method of grouping interventions [11]. Four intervention groups were defined for this study [13]
Intervention measure—Fidelity	Not available in dataset
Intervention measure—Quality	Use of Omaha System guidelines for evidence-based family home visiting documentation in EHR [9, 10] Psychometrics of the Omaha System intervention Scheme were examined during development [5]. It was field tested at the VNA of Omaha and test agencies in Des Moines, Delaware, and Indianapolis. The percentage of agreement between the staff nurse and nurse testers, and between testers was computed for each general intervention. Percentages of agreement ranged from 42.2 to 96.9% with eight of the twelve percentages at or above 80% [12]
Interventionist—Qualifications	The PHNs were bachelor of science or masters-prepared nurses [4]
Interventionist—Training	Agencies supported public health nurses by providing training in accordance with state and agency requirements for additional continuing education [7, 9]
Interventionist—Demographics	Not available in dataset
Interventionist—Organization	Not available in dataset
Source of data	An existing de-identified dataset generated by PHNs during routine documentation was reused

Worksheet D Define the Outcome and Related Measures or Scales

Definition of the Outcome: Maternal Risk Index.

Definition of the problem	Mothers with Mental health problems are at risk for and suffer disproportionately from poor health and social outcomes [1–4]
Measure/scale that operationalize the problem	Maternal Risk Index (MRI) based on number of Omaha System Problems (numerator) and baseline Knowledge scores for these problems (denominator) [6]
Use of measure in previous research	Previously the MRI has been shown to predict variability in health outcomes for high risk mothers [6]
Psychometric properties	The MRI is a metric that was developed using the Omaha System Problem Rating Scale for Outcomes [6]. The formula takes into account known maternal risk factors, including number of client Problems, baseline Knowledge scores, and co-morbidities and challenges such as poverty, mental health issues (Mental health problem), domestic violence (Abuse problem), and substance abuse (Substance use problem) [6]
Validity (construct)	Studies have shown positive associations between the MRI and duration of PHN services, indicating intervention tailoring for high risk mothers [6]

Reliability (internal consistency)	An existing dataset was reused
Reliability (test-retest)	An existing dataset was reused
Reliability (across raters)	An existing dataset was reused
Data collection—Strategy	Public health nurses documented problem assessments including Knowledge, Behavior, and Status scores in the electronic health record. Data were abstracted and de-identified for use in the study. An algorithm was used to transform raw data into MRI scores [6]
Data collection—Training	Public health nurses participated in regular training in the use of the Problem Rating Scale for Outcomes for documentation of Problem-specific assessments during routine team meetings [7, 9]
Data collection—Timing	The ratings were recorded when Problems were identified (assessed), and again when these Problems were re-evaluated or resolved [4]
Source of data	An existing dataset was reused.

Worksheet E Plan the Analysis Methods

Exploratory data analysis: sample characteristics	
Exploratory data analysis: interventions	Bar and line graphs to visualize intervention groups, defined by proportions of Omaha System intervention type
Exploratory data analysis: Outcomes	
Descriptive analysis: Sample	Standard descriptive statistics (mean, SD) to describe the groups of adults and adolescents with and without Mental health problems [4, 13]
Descriptive analysis: Interventions	Intervention proportions by Problem and Category overall, and for each intervention group
Descriptive analysis: problems	
Descriptive analysis: outcome P_{Time1}, P_{Time2}	Analysis of overall Maternal Risk Index relative to Problem Rating Scale for Outcomes scores for the total sample and by group
Inferential analysis: outcome P_{Time1}, P_{Time2}	Inferential statistics to test the significance of patterns relating client characteristics, Problems, and Categories for the four groups using one-way ANOVA tests and two-way ANOVA tests with interactions, and ANCOVA for testing individual knowledge, behavior, and status outcomes, adjusted by age and having the Mental health problem (yes/no) [13]
Inferential analysis: effect size	Cohen's d to evaluate effect size of outcomes overall and by group
Inferential analysis: benchmark attainment	
Inferential analysis: survival	Not applicable
Time analysis	Before and after intervention
Inferential analysis: correlations	No correlation analysis planned for this study
Qualitative analysis	Not applicable

Worksheet F Results and Alignment with the Literature

Sample characteristics	There were 676 mothers ages 14–52 in the sample. The proportion of adolescents was relatively higher in intervention group1 (n = 34) and intervention group4 (64), than in intervention group2 (16) and intervention group3 (14)
How sample aligns with literature or demographics of the population	This is consistent with the proportion of mothers with Mental health problems in previous research [1–3, 6, 7]
Interventions	There were 78,823 interventions, with an average of 116.6 per client. The number of interventions by intervention group differed ($t(675) = 3.88$, $p < 0.01$). The number of interventions for the Mental health problem differed by group ($t(675) = 4.42$, $p < 0.01$). The round-up round-down method was successful in creating four intervention groups with different characteristics and enabling knowledge discovery on differential intervention delivery and associated outcomes [11]
How interventions align with the literature	Intervention groups discovered in this study are novel and should be evaluated in other datasets [13]. Alternative explanation: The intervention groups may represent different programs that are offered through the public health department, and these programs may serve clients with differing risk levels, which would likely relate to final outcome attainment
Outcomes	After intervention, there was a positive change in overall outcomes for all groups (all $p < 0.01$), with variability between groups in each intervention group for Knowledge outcomes. Outcome variability was primarily related to age
How outcomes align with the literature	Outcome findings align with previous literature on PHN outcomes [6, 7, 14, 15], and expands the knowledge base regarding PHN interventions for adolescent and adult mothers with and without the Mental health problem
Relationship between interventions and outcomes	The best knowledge outcomes were for clients in intervention group 3 who received the most Teaching, Guidance, and Counseling compared to Surveillance, and a small proportion of Case Management [13]. Between group analysis showed significant interactions for only the knowledge outcomes. Adolescents in intervention group 1 had greater improvement in knowledge than adults in intervention groups 1, 2, and 4. Adults with the Mental health problem had lower Knowledge change compared to adults without the Mental health problem in intervention groups 1 and 2, and lower Knowledge change compared to adolescents in intervention group 4 [13]. Alternative explanation: The best outcomes were found in the intervention group with most Teaching, Guidance, and Counseling, less Surveillance, and a small amount of Case Management. This may be due to intervention tailoring for clients having a higher aptitude for learning who need less monitoring and fewer referrals
How intervention/outcome relationships align with the literature	Further study is needed regarding intervention group composition and outcomes, specifically proportionally higher Teaching, Guidance, and Counseling and low Case Management

References

1. World Health Organization (WHO) (2014) Social determinants of mental health. [Internet]. [cited 2017May12].Availablefrom:http://apps.who.int/iris/bitstream/10665/112828/1/9789241506809_eng.pdf
2. Rahman A, Surkan PJ, Cayetano CE, Rwagatare P, Dickson KE (2013) Grand challenges: integrating maternal mental health into maternal and child health programmes. PLoS Med 10(5):e1001442
3. Goetzel RZ, Long SR, Ozminkowski RJ, Hawkins K, Wang S, Lynch W (2004) Health, absence, disability, and presenteeism cost estimates of certain physical and mental health conditions affecting US employers. J Occup Environ Med 46(4):398–412
4. Garcia C, McNaughton D, Radosevich DM, Brandt J, Monsen K (2013) Family home visiting outcomes for Latina mothers with and without mental health problems. Public Health Nurs 30(5):429–438
5. Martin KS (2005) The Omaha System: a key to practice, documentation, and information management. Reprinted, 2nd edn. Health Connections Press, Omaha, NE
6. Monsen KA, Radosevich DM, Kerr MJ, Fulkerson JA (2011) Public health nurses tailor interventions for families at risk. Public Health Nurs 28(2):119–128
7. Monsen KA, Brandt JK, Brueshoff B, Chi CL, Mathiason MA, Swenson SM, Thorson DR (2016) Social determinants and health disparities associated with outcomes of women of childbearing age who receive public health nurse home visiting services. J Obstet Gynecol Neonatal Nurs 46(2):292–303
8. Barnard KE, Magyary D, Sumner G, Booth CL, Mitchell SK, Spieker S (1988) Prevention of parenting alterations for women with low social support. Psychiatry 51:248–253
9. Monsen KA, Swenson SM, Kerr MJ (2016) The perceptions of public health nurses on using standardized care plans to translate evidence-based guidelines into family home visiting practice. Kontakt 18:e75–e83
10. Omaha System Guidelines. Intervention library: family home visiting. [Internet]. [cited 2017 May 12]. Available from: http://omahasystemguidelines.org/?p=303
11. Chi C. (2015) Data mining for translation to practice. Presented at second international conference on research methods for standard terminologies, West Paul, April 15
12. Martin KS, Scheet NJ (eds) (1992) The Omaha System: applications for community health nursing. WB Saunders Company, Philadelphia, PA
13. Park YS (2017) Use of novel intervention management methods with large nursing datasets to model public health nurse home visiting intervention approaches and outcomes. Poster presented at the Midwest Nursing Research Society Annual Meeting, Apr 2017, Minneapolis, MN
14. Monsen KA, Chatterjee SB, Timm JE, Poulsen JK, McNaughton DB (2015) Factors explaining variability in health literacy outcomes of public health nursing clients. Public Health Nurs 32(2):94–100
15. Monsen KA, Radosevich DM, Johnson SC, Farri O, Kerr MJ, Geppert JS (2012) Benchmark attainment by maternal and child health clients across public health nursing agencies. Public Health Nurs 29(1):11–18
16. Nielsen K, Randall R (2013) Opening the black box: presenting a model for evaluating organizational-level interventions. Eur J Work Organ Psy 22(5):601–617
17. Drolet BC, Lorenzi NM (2011) Translational research: understanding the continuum from bench to bedside. Transl Res 157(1):1–5
18. Maternal infant and early childhood home visiting [US DHHS] (2016) [Internet]. [cited 2017 May 12]. Available from: http://mchb.hrsa.gov/programs/homevisiting/index.html
19. Nurse-Family Partnership (2014) [Internet]. [cited 2017 May 12]. Available from: https://www.nursefamilypartnership.org/
20. Byrne D (2013) Evaluating complex social interventions in a complex world. Evaluation 19(3):217–228

21. Petticrew M (2011) When are complex interventions 'complex'? When are simple interventions 'simple'? Eur J Public Health 21(4):397–398
22. Erickson KJ, Monsen KA, Attleson IS, Radosevich DM, Oftedahl G, Neely C, Thorson DR (2015) Translation of obesity practice guidelines: measurement and evaluation. Public Health Nurs 32(3):222–231
23. Silverman B, Mai C, Boulet S, O'Leary L. Logic models for planning and evaluation. The Division of Birth Defects and Developmental Disabilities (DBDDD), Centers for Disease Control and Prevention (CDC): Atlanta GA
24. Krueger RA. Evaluation. [Internet]. [cited 2017 May 12]. Available from: http://www.tc.umn.edu/~rkrueger/evaluation.html
25. Astleitner H. A systematic approach on the theoretical quality of educational intervention research: the Intervention Theory Questions (ITQ). Paper presented at European Conference on Educational Research (ECER), Istanbul, 11–15 Sept 2013
26. American Nurses Association [ANA] (2012) ANA recognized terminologies that support nursing practice [Internet], [cited 2017 May 12]. Available from: http://www.nursingworld.org/MainMenuCategories/Tools/Recognized-Nursing-Practice-Terminologies.pdf
27. Huber D, Delaney C (1997) The American Organization of Nurse Executives (AONE) research column. The nursing management minimum data set. Appl Nurs Res 10:164–165
28. Marucci AR, De Caro W, Petrucci C, Lancia L, Sansoni J (2015) ICNP-International Classification of Nursing Practice: origin, structure and development. Prof Inferm 68(2):131
29. Butcher HK, Bulechek GM, Dochterman JM, Wagner C (2013) Nursing interventions classification (NIC). Elsevier, Amsterdam
30. Clinical Care Classification [Internet] [cited 2017 May 12]. Available from: https://www.sabacare.com/about/
31. Monsen KA, Peterson JJ, Mathiason MA, Kim E, Lee S, Chi CL, Pieczkiewicz DS (2015) Data visualization techniques to showcase nursing care quality. Comput Inform Nurs 33(10):417–426
32. Monsen KA, Fitzsimmons LL, Lescenski BA, Lytton AB, Schwichtenberg LD, Martin KS (2006) A public health nursing informatics data-and-practice quality project. Comput Inform Nurs CIN 24(3):152–158
33. Camp RC, Tweet AG (1994) Benchmarking applied to health care. Jt Comm J Qual Improv 20(5):229–238
34. Glanz K, Bishop DB (2010) The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 31:399–418
35. National Institutes of Health Office of Behavioral and Social Science [US]. Social and behavioral theories. [Internet]. [cited 2017 May 12]. Available from: http://www.esourceresearch.org/eSourceBook/SocialandBehavioralTheories/1LearningObjectives/tabid/724/Default.aspx
36. Acton GJ, Irvin BL, Hopkins BA (1991) Theory-testing research: building the science. Adv Nurs Sci 14(1):52–61
37. Velicer WF, Prochaska JO (1999) An expert system intervention for smoking cessation. Patient Educ Couns 36(2):119–129
38. Pender NJ (2011) Heath promotion model manual. University of Michigan, Ann Arbor, MI
39. Monsen KA, Attleson IS, Erickson KJ, Neely C, Oftedahl G, Thorson DR (2015) Translation of obesity practice guidelines: interprofessional perspectives regarding the impact of public health nurse system-level intervention. Public Health Nurs 32(1):34–42
40. Santacroce SJ, Maccarelli LM, Grey M (2004) Intervention fidelity. Nurs Res 53(1):63–66
41. Nelson MC, Cordray DS, Hulleman CS, Darrow CL, Sommer EC (2012) A procedure for assessing intervention fidelity in experiments testing educational and behavioral interventions. J Behav Health Serv Res 39(4):374–396
42. Georgia State University Tobacco Portal. Databases and datasets. [Internet]. [cited 2017 May 12]. Available from: Available at: http://tobacco.publichealth.gsu.edu/resources/data/
43. Institute of Medicine [US] Committee on a National Surveillance System for Cardiovascular and Select Chronic Diseases (2011) A nationwide framework for surveillance of cardiovascular and chronic lung diseases. 5. Existing surveillance data sources and systems. National Academies Press, Washington, DC

Define the Outcome (Worksheet D)

13

13.1 Preparing to Complete Worksheet D

In intervention effectiveness research, quality improvement activities, and program evaluation, **defining and measuring the problem** enables the comparison of problem changes over time. In this chapter a second fictitious study is introduced: changes over time in health outcomes of individuals receiving brief family home visiting services compared to those receiving long term visits, and specific instructions for completing Worksheet D are provided with references related to the study [1–18]. From the population health perspective, positive change in one or more health status indicator(s) is the ultimate goal of all interventions [19] Outcome Indicator(s) pertain to the problem, translating concepts into specific measures that can be collected, analyzed, and interpreted. Specifically in the PIO MM, outcome indicators are valid, reliable measures used to monitor and collect meaningful evidence about a problem over time [20]. Consider the problem defined in Worksheet B, and think about one or more meaningful measures to monitor and collect data regarding the problem. It is preferable to measure more than one outcome indicator for multiple dimensions of the problem in order to compare the change(s) across measures. Well-defined outcome indicators are a precondition for rigorous intervention effectiveness research, quality improvement activities, and program evaluation [20].

Data quality begins with the selection of the measure(s) and also includes rigorous data collection procedures, data management, and training as described in the preceding chapters [20]. Quality refers to the integrity of data that are used in the study/project. High-quality data are reliable, valid, and suitable for use in the study [20]. Observational data are known to have limitations that should be accounted for in study design and addressed in the interpretation of findings [21–22]. Obtaining existing data for PIO MM studies depends on the availability of high quality data to operationalize PIO MM concepts. Such data are becoming more readily available as routine use of EHRs for clinical documentation increases; however, documenting quality of EHR data is uncommon. Therefore it is critical to consider the potential

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_13

sources of data and select those that offer best quality while operationalizing PIO MM concepts [20].

In intervention effectiveness research, quality improvement activities, and program evaluation, the outcome is measured immediately following the intervention, and is a measure of the problem (P) over time (P_{Time1}, P_{Time2}, … P_{TimeX}). Most commonly, studies evaluate the difference between problem-specific measures for participants at two points in time, before and after intervention (P_{Time2}–P_{Time1}). Thus, it is ideal to operationalize problem using measures and/or scales. The problem defines the outcome, and operationalizing the outcome using valid, reliable measure(s) that are specific to the problem enables robust outcome measurement [22]. For example, for intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?* P_{Time2} = equity measure X on discharge P_{Time1} = health equity measure X on admission.

For quality improvement: *was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* P_{Time2} = cigarette smoking measure X on discharge P_{Time1} = cigarette smoking measure X on admission. For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* P_{Time2} = obesity measure X on discharge P_{Time1} = obesity measure X on admission. The corollary questions for benchmarking of outcomes compare P_{Time2} scores (benchmark attainment at discharge or a specified point in time).

13.2 Step-by-Step Instructions for Completing Worksheet D

Given that there may be more than one measure of the outcome, Worksheet D should be completed for each outcome measure that will be used in your project. At the top of Worksheet D, state the instrument, scale, or measure that will be described in the worksheet. In the first row, **restate the definition of the problem** (Worksheet B) to ensure consistency between the problem and the measure(s) used to operationalize the problem, and then note the reason for selecting measure that will be used to **operationalize the problem**, with reference. Note **use of measure in previous research** as described in the literature. This information should substantiate the alignment of the problem and its measure. Measurement of health equity requires three components, a measure of health status, a group comparison measure, and a method for comparing health status across groups [23]. Therefore to measure health equity both health status and demographic characteristics are needed. It may be difficult to obtain objective smoking behavior assessments. A self-report of smoking has been studied and found to be valid and reliable in an anonymous on-line survey [24] which may be of value in the proposed smoking cessation evaluation. For obesity among individuals with diabetes, BMI may not be the best or most relevant measure of obesity [25, 26]. Other measures such as waist circumference, waist-to-hip ratio, skinfold thicknesses, and bioelectrical impedance may be of value with BMI depending on other factors [27]. For the fictitious study of home visiting

program outcomes of individuals receiving brief family home visiting services compared to those receiving long term visits, Worksheet D has been completed for the Omaha System Problem Rating Scale for Outcomes, which operationalizes three dimensions of problem-specific outcomes: Knowledge, Behavior, and Status [8].

Next, describe the psychometric properties of the measure. The next several prompts are provided to assist in describing the psychometric properties of the measures based on the literature, including **construct validity** and three types of reliability: **internal consistency, test-retest,** and **reliability across raters**. Psychometric properties may be described in seminal publications about measure development or may be reported in studies that used the measure. Whenever possible refer to original sources for detailed information regarding the types of psychometric tests that were performed, and findings that support use of the measure; as well as confidence in the data that will be (or were) generated by the measure. For the fictitious study of home visiting program outcomes of individuals receiving brief family home visiting services compared to those receiving long term visits, detailed descriptions of reported psychometrics of the Omaha System Problem Rating Scale for Outcomes are described [17].

Next, describe data collection procedures including strategy, training, and timing relative to the intervention. For secondary analyses of existing data, it is essential to understand the primary data collection procedures as well as the way in which the existing dataset was obtained. For example, if using the anonymous on-line survey for self-report of smoking, it would be important to provide information about how the data will be obtained and matched before and after intervention. For obesity among individuals with diabetes, information about how BMI (or other) measurement will be (or were) obtained will be critical for interpreting the results and replicating the study. For the fictitious study of home visiting program outcomes of individuals receiving brief family home visiting services compared to those receiving long term visits, details regarding obtaining the datasets A, B, and C are given, together with the information that all data were obtained from individuals who consented that their data could be used in research.

Finally describe the source the outcome data, and check to make sure it aligns with the source of intervention data identified in Worksheet C. For example, if you are planning to use a publicly available dataset of interventions related to smoking cessation or obesity that is available for the population of interest, ensure that the dataset includes all the variables that are needed based on PIO MM [188–189]. Likewise, if your project will take place within a health care organization there may be existing EHR data that meets the requirements of the intervention variable but not the outcome measure, or vice versa. As noted previously, all measures selected for PIO MM-based studies should be chosen based on their ability to meaningfully operationalize the PIO MM concepts. For the fictitious study of home visiting program outcomes of individuals receiving brief family home visiting services compared to those receiving long term visits the sources of data were noted in the data collection strategy section.

The problem measure must be understood in context. Measuring program performance by tracking outcome indicator(s) in a single group without other measures may lead to improved documentation of the problem in a particular population, practice, or setting, despite the intervention strategies that may be addressing the

problem. Thus, tracking only the presence or absence of a problem may result in misleading findings of increased problem incidence and prevalence when in fact the problem may have been present but previously undetected [20]. Multiple measures may address this threat to the validity of the findings. To achieve high quality, reproducible results, it is essential to adhere to rigorous methods, combining multiple well documented outcome indicators with contextual information including sample characteristics, setting characteristics, and intervention descriptions as shown in the PIO MM [20]. Upon completion of worksheet D, the reader will have the necessary foundation for finalizing study purpose, design, and analysis approach as described in Chapters 4–7. A completed PIO MM diagram for this example shows relationships among the PIO MM concepts (Fig. 13.1).

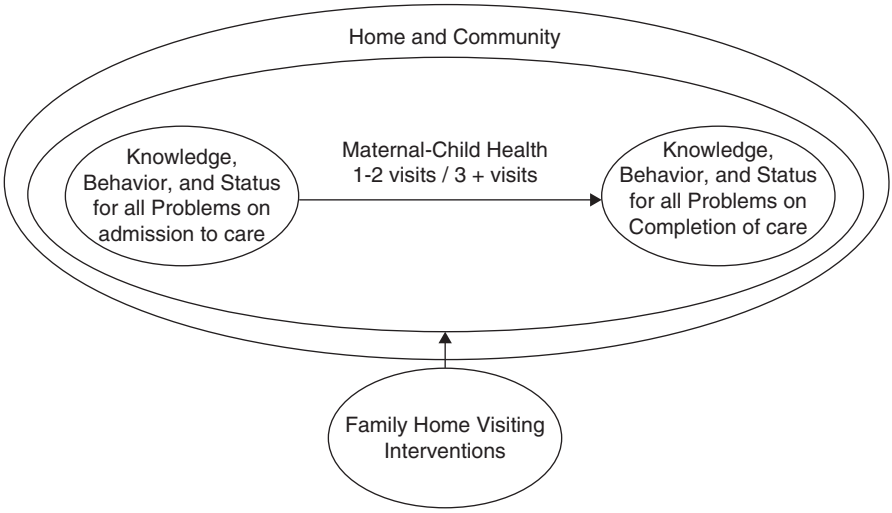

Fig. 13.1 PIO MM for evaluation of health outcomes of individuals receiving 1–2 maternal-child public health nursing home visits compared to individuals receiving three or more visits. This example is loosely based on a study by Kathy Dubbels, Jill E. Timm, and Karen A. Monsen. The results are fictional and are intended to illustrate the types of data analyses that explore and demonstrate care quality and outcomes for groups within an existing dataset

Reflection Questions

- What is the value and/or importance of using valid, reliable measures?
- How does using existing data threaten the validity of the project? Improve the feasibility of the project?
- In the ideal intervention effectiveness research or program evaluation project, what measures should be used to operationalize the outcome over time? Is this possible? If not, what alternatives are available?
- How can reliability be assured in intervention effectiveness research or program evaluation?

Worksheet A. References from the Preliminary Literature Review

- Institute of Medicine (US). Committee for the Study of the Future of Public Health. The future of public health. National Academy Press; 1988 [1].
  - Describes three core functions and ten essential services of public health
  - Establishes a foundation for understanding various reasons for providing short or long-term visits
- Public Health Observer. Public health surveillance systems. http://publichealthobserver.com/public-health-surveillance-systems/ [2]
  - Clarifies the role of surveillance in public health programs
- United States Department of Health and Human Services, Health Resources and Services Administration [US DHHS HRSA]. (2016). Maternal infant and early childhood home visiting. Retrieved from http://mchb.hrsa.gov/programs/homevisiting/index.html [3]
  - Describes evidence-based maternal, infant, and early childhood home visiting programs

Worksheet B. Definition of the Problem with Statement of the Gap in Knowledge

Problem: Variability in Outcomes of Public Health Nurse Home Visiting Programs.

Definition of the problem	The goal of public health is to improve the health status of the population, with careful attention to and respect for the perspectives and values of the diverse members of the community being served [1, 3]. Public health nurse home visits are a strategy to improve maternal and child health, as well as to prevent child abuse and neglect, encourage positive parenting, and promote child development and school readiness [3]
Population of interest	Individuals in the population of mothers and children who are at risk of poor health outcomes
Background literature describing the problem (% or number of individuals in the total population)	Nearly half (41%) of the nation's childbearing families are at high risk [4, 5]. In 2009, 31% of children lived in low-income families, defined as income below 200% of the federal poverty level, with fewer in urban (29%) vs. rural (37%) and native born (28%) vs. immigrant (65%). Only 23% of white children live in low-income families, vs. black (72%), Hispanic (66%), American Indian (63%), and Asian (39%) [5]
Background (costs incurred addressing the problem)	Excess costs incurred by society for high risk families were estimated at $160 million per year in one jurisdiction alone [6]. The average annual cost of adolescent childbearing was estimated at $7.3 billion in 2008 [7]
Background (years of potential life lost due to the problem)	
Background (health system-related gaps in addressing the problem)	There is compelling evidence of failure to prevent costly health and social problems for mothers and children in high risk families, many of whom suffer from disparities in preterm birth, child abuse, school failure, and welfare dependence [4]

Problem measurement instrument/scale	Variable health outcomes of mothers and children receiving public health nurse home visiting interventions as measured by Omaha System Problem Rating Scale for Outcomes on admission and dismissal from Home Care services [8]
Anticipated outcome and rationale	It is likely that family home visiting interventions interventions improve health outcomes of high risk mothers and children [6–7]. Previous studies have shown variability in health outcomes across populations and programs [3, 9–11]
What is not known?	Home visiting programs differ by number of visits. Most family home visiting programs consist of many visits. It is not known whether improvement in the health of those mothers and children who receive only 1–2 visits, compared to 3+ visits. It is not known if programs that provide brief home visits to high risk mothers and children are useful to individuals and affect the health of populations

Worksheet C. Describe the Intervention Used to Address the Problem

Definition of the Intervention	The intervention is defined as a 'visit' in this study, and groups based on the number of visits (1–2 vs. 3+) were created based on the number of visits
Percentage of persons who improved after intervention	The percentages of individuals who improved after intervention varied by problem and there was a ceiling effect on improvement for problems that were rated higher on admission to care [9]
Meta-analysis showing levels of effectiveness across studies	Extensive literature on home visiting shows levels of effectiveness across studies but no information is available for programs that provided brief home visits [3]
For whom	A randomized comparison of home visits and hospital-based group follow-up visits after early postpartum discharge showed improved outcomes for mothers receiving brief home visits [12]
Under what conditions	Early postpartum discharge [12]
Theory of causal mechanism	Clinical Nursing Models middle range theory or theoretical framework derived from a randomized trial of maternal-child intervention effectiveness [13]
Essential (core) components	Evidence-based family home visiting guidelines [14–16]
Intervention content	Not available in the dataset
Intervention adherence and involvement	Not available in the dataset
Intervention measure—amount	Number of visits
Intervention measure—type	Not available in the dataset
Intervention measure—fidelity	Not available in the dataset
Intervention measure—quality	Not available in the dataset
Interventionist—qualifications	Public health nurses were registered nurses (RNs), usually bachelor's or master's prepared
Interventionist—training	Public health nurses with additional training and experience in home visiting

Interventionist—demographics	Not available in the dataset
Interventionist—organization	Not available in the dataset
Source of data	Re-use of data from three sources, including (A) data voluntarily contributed by county PHN agencies for a statewide program evaluation for a 1 year period (2007), (B) data aggregated by a national software vendor from multiple PHN agencies and used with agency permission (2005–2011), and (C) secondary use of a large practice-generated research data set from one PHN agency (2000–2005)

Worksheet D. Define the Outcome and Related Measures or Scales

Definition of the problem	Poor health outcomes of mothers and children. This study examines a large dataset of mothers and children receiving brief vs. long term home visits from public health nurses. The goal of public health is to improve the health status of the population, with careful attention to and respect for the perspectives and values of the diverse members of the community being served [1, 3]. Public health nurse home visits are a strategy to improve maternal and child health, as well as to prevent child abuse and neglect, encourage positive parenting, and promote child development and school readiness [3]
Measure/scale that operationalize the problem	Omaha System Problem Rating Scale for Outcomes: Knowledge, Behavior, and Status Scales [8]. The Problem Rating Scale for Outcomes consists of three five-point Likert-type ordinal rating scales; one each for the concepts of Knowledge, Behavior, and Status (KBS). It is used as an assessment relative to all Omaha System problems. The scoring of the scales ranges from 1 (most negative) to 5 (most positive). For Knowledge, 1 = No knowledge and 5 = Superior knowledge. For Behavior, 1 = Not appropriate behavior and 5 = Consistently appropriate behavior. For Status, 1 = Extreme signs/symptoms and 5 = No signs/symptoms
Use of measure in previous research	Extensive use in previous intervention effectiveness research, quality improvement activities, and program evaluation, especially in public health settings [8–11]
Psychometric properties	Problem Rating Scale for Outcomes is a valid, reliable measure with psychometric properties established during development and subsequently during application in practice and research [8, 17]
Validity (construct)	Content validity of the PRSO was assessed using a panel-of-experts approach. Ten Omaha System problems were randomly selected for examination. Experts were recruited based on their credentials and expertise in practice related to representative Omaha System content. The experts reviewed problem definition, KBS subscales, a copy of the PRSO, and prototypical guidelines for each of the selected problems. They rated each item from 1 (definitely no) to 4 (definitely yes) using the Content Validity Index (CVI). Analysis of the results was accomplished through calculation of the proportion of times the experts chose a rating of 3 or 4 for an item. The composite proportion for knowledge was 0.85, for behavior was 0.81, and for status was 0.77. When CVI values for specific units were averaged, any prototype statements or sub scale items below 0.80 were revised before they were published [17]

Reliability (internal consistency)	
Reliability (test-retest)	
Reliability (across raters)	Reliability of the Problem Rating Scale for Outcomes KBS scales was studied using a research assistant who accompanied nurses on 97 visits and compared independent KBS ratings following the visits. The research assistant and nurse ratings were analyzed for agreement using a coefficient gamma test, and were found to agree significantly ($p < 0.01$). Coefficient gamma for knowledge ratings was 0.53, for behavior ratings was 0.60, and for status ratings was 0.87. The percentage of agreement between staff nurse and raters and between raters was computed for exact matches and for differences of one. Exact matches ranged from 11.7 to 64.8%, and differences of one ranged from 82.6 to 96.1% [17]
Data collection—strategy	Public health nurses assessed and documented the KBS ratings during routine care, resulting in the existing datsets (A, B, C) that were provided with consent of the individuals for use in research. By source, (A) data voluntarily contributed by county PHN agencies for a statewide program evaluation for a 1 year period (2007), (B) data aggregated by a national software vendor from multiple PHN agencies and used with agency permission (2005–2011), and (C) secondary use of a large practice-generated research data set from one PHN agency (2000–2005)
Data collection—training	Not available in the dataset
Data collection—timing	Admission to care and dismissal from care
Source of data	Re-use of data from three sources (A, B, C)

Worksheet E. Plan the Analysis Methods

Exploratory data analysis: Sample characteristics	No exploratory data analysis planned
Exploratory data analysis: Interventions	Not available in the dataset
Exploratory data analysis: Outcomes	Patterns in outcomes by brief vs. continuing visits groups (1–2/3+) and sample (A, B, C) using line graphs and/or heat maps
Descriptive analysis: Sample	Number of cases in each sample (A, B, C). All mothers and children qualified for public health nurse home visits (low income, high risk families)
Descriptive analysis: Interventions	Not available in the dataset
Descriptive analysis: Problems	Not available in the dataset but have been described in the literature [9–11, 14, 15]
Descriptive analysis: Outcome P_{Time1}, P_{Time2}	Descriptive analysis of KBS outcomes by brief vs. continuing visits groups (1–2/3+) and data source (A, B, C) using means, SD, frequencies
Inferential analysis: Outcome P_{Time1}, P_{Time2}	Significance of KBS outcomes (P_{Time2}-P_{Time1}) and by group and sample using parametric or nonparametric tests
Inferential analysis: Effect size	Effect size of overall outcomes using Cohen's *d*

Inferential analysis: Benchmark attainment	Significance of KBS change outcomes (P_{Time2}-PT_{ime1}) and by group using parametric or nonparametric tests
Inferential analysis: Survival	No survival analysis planned
Time analysis	Time is included within the groups as number of visits
Inferential analysis: Correlations	No correlations analyses planned
Qualitative analysis	Not available in dataset

Worksheet F. Results Statements and Alignment with the Literature

Sample characteristics	There were 25,855 individuals in the sample, with only 799 (3%) receiving 1–2 visits. All qualified for public health nurse home visits as specified by the Title V Maternal and Child Health Services Block Grant Program (i.e. low income, high risk families) [18]
How sample aligns with literature or demographics of the population	Not available in dataset, but it is expected to align with other home visiting samples described in the literature because of the consistent compliance with risk criteria used for admission to public health nursing services [9, 15]
Interventions	Evidence-based family home visiting interventions have been described in the literature [9–11, 14–15]
How interventions align with the literature	Not available in dataset, but it is expected to align with other home visiting samples described in the literature because of the consistent emphasis on evidence-based home visiting practice in public health nursing [9, 15]
Outcomes	EDA revealed patterns in descriptive outcomes for KBS on admission and discharge by number of visits (1–2/3+) and by data source (A, B, C). Pattern 1: On average, KBS scores for all groups improved after receiving visits. This pattern was statistically significant with a medium effect size across all samples: knowledge ($t(25855) = 7.43$, $p < 0.01$, $d = 0.5$), behavior ($t(25855) = 6.36$, $p < 0.001$, $d = 0.4$), and status ($t(25855) = 5.52$, $p < 0.01$, $d = 0.44$). Pattern 2: All admission scores for individuals receiving 1–2 visits were higher than those for individuals receiving 3+ visits across all datasets. This pattern was significant for all datasets, with greater effect sizes for the 3+ visits group. Pattern 3: For all change scores except the difference in status for data source A, individuals receiving 3+ visits improved more than those who received 1–2 visits
How Outcomes align with the literature	This aligns with previous literature regarding effectiveness of public health nurse home visiting [3, 9–11, 12, 15] and adds to what is known about the outcomes of individuals that received 1–2 visits. This is consistent with the notion that public health nurses tailor interventions to meet unique individual needs [11]. Findings are novel and should be replicated with other family home visiting data
Relationship between interventions and outcomes	All three data sources (A, B, C) showed the same pattern of outcome for each of the visited conditions (1–2/3+)
How Intervention/Outcome Relationships align with the literature	This aligns with previous literature regarding use of large datasets to replicate analyses in order to increase confidence in the internal validity of large datasets [19]

References

1. Institute of Medicine [US] (1988) Committee for the study of the future of public health. In: The future of public health. National Academy Press, Washington, DC
2. Public Health Observer. Public health surveillance systems. [Internet]. [cited 2017 May 12]. Available from: http://publichealthobserver.com/public-health-surveillance-systems/
3. Maternal infant and early childhood home visiting [US DHHS], 2016. [Internet]. [cited 2017 May 12]. Available from: http://mchb.hrsa.gov/programs/homevisiting/index.html
4. Hoffman SD, Maynard RA (2008) Kids having kids: the economic costs and social consequences of adolescent pregnancy. The Urban Institute Press, Washington, DC
5. National Center for Children in Poverty, Mailman School of Public Health, Columbia University, [Internet]. [cited 2017 May 12]. Available from: http://nccp.org/profiles/state_profile.php?state=MN&id=6.
6. Roth J, Figlio DN, Chen Y, Ariet M, Carter RL, Resnick MB, Morse SB (2004) Maternal and infant factors associated with excess kindergarten costs. Pediatrics 114(3):720–728
7. Larson K, Halfon N (2010) Family income gradients in the health and health care access of US children. Matern Child Health J 14:332–342
8. Martin KS (2005) The Omaha System: a key to practice, documentation, and information management, 2nd edn. Health Connections Press, Omaha, NE
9. Monsen KA, Fulkerson JA, Lytton AB, Taft LL, Schwichtenberg LD, Martin KS (2010) Comparing maternal child health problems and outcomes across public health nursing agencies. Maternal Child Health J 14(3):412–421
10. Monsen K, Sanders A, Yu F, Radosevich D, Geppert J (2011) Family home visiting outcomes for mothers with and without intellectual disabilities. J Intellect Disabil Res 55(5):484–499
11. Monsen KA, Radosevich DM, Kerr MJ, Fulkerson JA (2011) Public health nurses tailor interventions for families at risk. Public Health Nurs 28(2):119–128
12. Escobar GJ, Braveman PA, Ackerson L, Odouli R, Coleman-Phox K, Capra AM, Wong C, Lieu TA (2001) A randomized comparison of home visits and hospital-based group follow-up visits after early postpartum discharge. Pediatrics 108(3):719–727
13. Barnard KE, Magyary D, Sumner G, Booth CL, Mitchell SK, Spieker S (1988) Prevention of parenting alterations for women with low social support. Psychiatry 51:248–253
14. Monsen KA, Swenson SM, Kerr MJ (2016) The perceptions of public health nurses on using standardized care plans to translate evidence-based guidelines into family home visiting practice. Kontakt 18(2):e75–e83
15. Monsen KA, Swenson SM, Klotzbach L, Mathiason MA, Johnson KE (2017) Empirical evaluation of change in public health nursing practice after implementation of an evidence-based family home visiting guideline. Kontakt 19(3) e75–e85
16. Omaha System Guidelines. Intervention library: family home visiting. [Internet]. [cited 2017 May 12]. Available from: http://omahasystemguidelines.org/?p=303
17. Health Resources and Services Administration [US]. Title V Maternal and Child Health Services Block Grant Program. [Internet]. [cited 2017 May 12]. Available from: https://mchb.hrsa.gov/maternal-child-health-initiatives/title-v-maternal-and-child-health-services-block-grant-program
18. Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD (2001) Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 54(8):774–781
19. Swider SM (2002) Outcome effectiveness of community health workers: an integrative literature review. Public Health Nurs 19(1):11–20
20. Centers for Disease Control and Prevention [US] (1999) Framework for program evaluation in public health. MMWR 48(RR11):1–40
21. Dunn OJ, Clark VA (2009) Basic statistics: a primer for the biomedical sciences. Wiley, Hoboken, NJ
22. Streiner DL, Norman GR, Cairney J (2014) Health measurement scales: a practical guide to their development and use. Oxford University Press, New York

23. Centers for Disease Control and Prevention [US]. Healthy People 2020. Measurement of health equity and health disparities. [Internet]. [cited 2017 May 12]. Available from: http://www.healthypeople.gov/2010/hp2020/advisory/PhaseI/appendix11.htm
24. Ramo DE, Hall SM, Prochaska JJ (2011) Reliability and validity of self-reported smoking in an anonymous online survey with young adults. Health Psychol 30(6):693–701. doi:10.1037/a0023443.
25. Dixon JB, Egger GJ, Finkelstein EA, Kral JG, Lambert GW (2015) 'Obesity paradox' misunderstands the biology of optimal weight throughout the life cycle. Int J Obes 39:82–84
26. Rothman KJ (2008) BMI-related errors in the measurement of obesity. Int J Obes 32(Suppl 3):S56–S59
27. Harvard School of Public Health Obesity Prevention Source. Obesity measurement. [Internet]. [cited 2017 May 12]. Available from: https://www.hsph.harvard.edu/obesity-prevention-source/obesity-definition/how-to-measure-body-fatness

Plan the Analysis (Worksheet E)

14

14.1 Preparing to Complete Worksheet E

To ensure success, techniques for analyzing, synthesizing, and interpreting findings should be decided before beginning the project [1]. This chapter builds on the preceding worksheets and culminates in planning the analysis. Completing Worksheet A was a discovery process in which a case was identified for initiating a PIO MM study. Worksheet B substantiated the importance of the problem that will be addressed in the project. Completing C and D were steps toward operationalizing PIO MM concepts of intervention and outcome. In this chapter, project planning will be facilitated by completing Worksheet E. At this critical juncture, reviewing key points in progress to date will aid in ensuring that the plans for data analysis align with all previous work. It is essential that all aspects of the project plan align and are internally consistent, and therefore a review process is presented to guide preparation, followed by guidance on planning the analysis that reflects the study purpose, design, and variables.

There are six steps in the planning process for the implementation of intervention effectiveness research, quality improvement activities, and program evaluation projects:

1. Review project statements.
2. Select statements that are most applicable to the project and discipline.
3. Review design options.
4. State the design.
5. Review variables.
6. Plan for creating new variables.

Following are more details about each of these steps, including examples.

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_14

14.1.1 Step 1. Review Project Statements

Review the PIO MM Statements for the Gap in Knowledge, and Project Purpose, Hypothesis, Question, and Goal as described in Chapter 2 and provided here with examples. Keep in mind that some of these statements will apply to your project while others will not, depending on the perspective of the project (intervention effectiveness research, quality improvement, or program evaluation).

- **Statement of the Gap in Knowledge:** It is not known if the intervention addressing the problem that is (efficacious or effective) for a given population in (a controlled environment or with other populations in real world settings) also will be associated with positive outcomes in the identified setting and population. This is because there are not studies in the literature specific to the gap we have defined. As described in Chapter 11, examples related to the question for intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?* The gap in knowledge may be stated as: Little is known about differences in health equity after public health nursing interventions. For quality improvement: *was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* The gap in knowledge may be stated as: It is not known to what extent people in who smoked cigarettes on admission to care changed cigarette smoking behavior. For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* The gap in knowledge may be stated as: It is not known to what extent obesity decreased among people with diabetes who participated in the program. For the fictitious study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions, the gap in knowledge may be stated as: little is known about variability associated with frailty and social determinants in the health outcomes of older adults receiving home care services.
- **Project Purpose:** The purpose of this project is to examine the outcome of the problem after intervention in this setting and population. For example, for intervention effectiveness research: *the purpose of this project was to examine health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions*. For quality improvement: *the objective of this QI initiative was to study change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle*. For program evaluation: *The purpose of this project was to evaluate change in obesity among people with diabetes who participated in the BMI reduction program*. For the fictitious study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions: the purpose of this project was to examine variability associated with frailty and social determinants in the health outcomes of older adults receiving home care services. In particular, these templates for project statements may be useful for framing the project without going beyond the PIO MM single-group before-and-after design.

- **Project Hypothesis**: Interventions are associated with changes in problems (P_{Time2}-P_{Time1}) for the identified population in the described setting over the given timeframe. For example, for intervention effectiveness research: *public health nurse interventions were associated with reduced health equity (reduced disparities in health outcomes)*. For quality improvement: *The QI intervention was associated with a change in behavior of people who smoked cigarettes on admission*. For program evaluation: *the program was associated with reduced BMI for people with diabetes*. A null hypothesis would state these examples in the negative, for example, for intervention effectiveness research: *public health nurse interventions were not associated with reduced health equity (reduced disparities in health outcomes)*. For quality improvement: *The QI intervention was not associated with a change in behavior of people who smoked cigarettes on admission*. For program evaluation: *the program was not associated with changes in BMI for people with diabetes*. For the fictitious study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions: frailty and social determinants of health were not associated with variability in health outcomes of older adults receiving home care services.
- **Project Question**: Are interventions associated with changes in problems (P_{Time2}-P_{Time1}) for the identified population in the described setting over the given timeframe? For example, for the fictitious intervention effectiveness research project: are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?* For the fictitious quality improvement activity: is the QI intervention *associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* For the fictitious program evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program? For the fictitious study of intervention patterns associated with outcomes of adult and adolescent mothers with and without the mental health problem, are public health nurse home visiting interventions differentially associated with outcomes of adolescent and adult mothers with and without the mental health problem?
- **Project Goal**: Was the intervention associated with changes in problems (P_{Time2}-P_{Time1}) for the identified population in the described setting over the given timeframe? Specifying the problem intervention, setting and timeframe, and operationalizing these as measures support the development of SMART goals: Specific, Measurable, Achievable, Relevant and Time-framed [2]. SMART goals extend the above statements by adding numeric benchmarks including percentages of the sample that will improve related to the dosage of the interventions over a specified length of time. SMART goals may be developed for any project. For example for intervention effectiveness research: *75% of high risk mothers will show increased health equity (reduced disparities) compared to a matched sample of low risk mothers, after 12 months of public health nursing home visiting services in which they received at least 1 visit per month*. For quality

improvement: *75% of people who smoked cigarettes on admission to care will show a 10% reduction in the number of cigarettes smoked per day after receiving advice at least one time through our QI PDSA cycle*. For program evaluation: *75% of people with diabetes and obesity after participating in at least 80% of the program sessions will show at least 10% reduction in BMI*. These smart goals offer attainable benchmarks for the number of participants and also the level of outcome, while explaining the necessary dose of interventions that are thought to be needed to achieve the goal. For the fictitious study of home visiting program outcomes of individuals receiving brief family home visiting services compared to those receiving long term visits, a SMART goal for a public health agency may be suggested: *75% of high risk mothers will show increased knowledge for at least one problem after 1–2 public health nurse home visits*.

14.1.2 Step 2. Select Statements That Are Most Applicable to the Project and Discipline

Select from these project statements as appropriate, and develop the statements for your project. Selections will be based on the perspective of your discipline or setting (e.g. intervention effectiveness research in a doctoral program, quality improvement in a hospital intensive care unit, and program evaluation in a public health department).

14.1.3 Step 3. Review Design Options

Review Chapter 3 for clarification regarding the single-group before-and-after design as it relates to the project (prospective or retrospective, comparative or not comparative, observational). If using an existing dataset, it is a retrospective study. If the data were generated by clinicians, it is an observational study. If there are comparisons among groups within the sample, it is a comparative study. If all three are true, it is a retrospective, observational, comparative study. Conversely, if the project will use new data, it is a prospective study. If the data are provided by self-report from the participants, the project is observational if there was no randomization.

14.1.4 Step 4. State the Design

Develop a statement that describes the design of your project. For example, For intervention effectiveness research evaluating differences in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions, the design may reuse existing public health nursing documentation data, comparing groups within the sample. Therefore is a retrospective, observational, comparative study. For quality improvement studying the difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle, the design may include data from previous patients, as well as collected

prospectively for every patient who received the novel intervention. This QI PDSA cycle would be an observational, comparative study. For program evaluation, difference in obesity among people with diabetes who participated in the BMI reduction program may be a prospective study using questionnaires administered before and after the program. The fictitious studies described in Chapters 11–13 are all retrospective, observational, comparative studies that reused existing data to examine outcomes by groups formed based on demographic or intervention variables in the datasets.

14.1.5 Step 5. Review Variables

Review necessary variables. Worksheets A–D describe in detail the variables needed to operationalize the PIO MM in the project. To enable meaningful analysis for a particular population or setting, additional context variables will be needed. Create lists of context variables that describe the identified population (e.g. age, race/ethnicity, gender, signs/symptoms) and setting (geographic location, health system unit or type). Whenever possible these factors should be described as context for the interpretation of the findings, and should be presented by comparison groups in order to understand potential differences between groups that may relate to findings.

14.1.6 Step 6. Plan for Creating New Variables

Plan to create additional variables as necessary. If you need to create or transform variables from existing variables, describe how you will do so (clustering, risk scores). This may be as simple as transforming a continuous variable to a categorical variable for use in a particular EDA technique or statistical model, or it may be as complex as developing a new metric from existing data. For intervention effectiveness research: *was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?* Mothers may be stratified by a risk index variable created from existing variables. For quality improvement: *was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle*? A categorical smoking variable may be created from continuous variable (e.g. count of number of cigarettes smoked in a day). For program evaluation: *Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?* A variable may be created from two dates to show the length of time in the program. For example, for the fictitious study of home visiting program outcomes of individuals receiving brief family home visiting services compared to those receiving long term visits, the sample was divided into two groups by the number of visits for each individual (1–2 visits group and 3 or more visits group). Check to make sure you and your team will be able to accomplish the transformation within the allotted time. Advanced transformations such as clustering should not be attempted if time is limited (See Chapter 5).

14.2 Step-by-Step Instructions for Completing Worksheet E

Based on the project purpose or goal and the variables that you will use to operationalize the PIO MM concepts, choose appropriate methods in order to complete Worksheet E. The following options for planning the EDA and the analysis of sample, interventions, outcomes, multiple variables are explained in greater detail in Chapters 5–7.

14.2.1 Exploratory Data Analysis

Consider options for visual displays that may reveal patterns, generating hypotheses that can be evaluated statistically and lead to discovery of new information. Using the matrix provided in Chapter 7 to arrange PIO MM variables in comparison with other PIO MM variables for your project. Methods for EDA may include:

- Heat maps to discover patterns among categorical variables [3–6]
- Line graphs to discover patterns in continuous data [4, 7–8]
- Maps to discover patterns in data by geographic location [9]

For example, for the fictitious intervention effectiveness research project: are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?* Heat maps of number of interventions by age and outcome may show patterns related to subgroups defined by age and outcome. For the fictitious quality improvement activity: is the QI intervention *associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* Line graphs of the numbers of visits and changes in smoking behaviors by individual characteristics may reveal patterns that are specific to behavior change for certain groups. For the fictitious program evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program? Heat maps comparing final scores by gender and age may reveal patterns. For the fictitious study of home visiting program outcomes of individuals receiving brief family home visiting services compared to those receiving long term visits, line graphs of baseline and final values for all outcome measures by data source may reveal similarities and differences across data sources. As EDA is an iterative process of discover, it should not be limited to the planned techniques; rather, the exploration of the data may be guided by the plan and as discoveries emerge they may lead to further exploration using additional techniques [10–14].

14.2.2 Sample

Consider options for analysis of the sample as described in Chapters 5 and 6, to provide an understanding of the population for which the results may be interpreted. To describe the sample:

- Use frequencies to describe sample characteristics operationalized by categorical variables (number, percent) [15, 16]
- Use mean and standard deviation calculations for sample characteristics operationalized by continuous variables (mean, SD) [15, 16]
- Use a parametric test to evaluate the significance of differences between groups when the data meets the assumptions for each test (e.g. normality, homogeneity, independence) (*p-value*) [15, 16]
- Use a non-parametric test to evaluate the significance of differences between groups when the data does not meet the assumptions for an typical parametric test (*p-value*) [15, 16]
- Use frequencies to describe the distribution of the problem(s) within the sample (number, percent) [15, 16]

For example, for the fictitious intervention effectiveness research project: are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?* Describe the demographic characteristics of the high risk mothers in the sample by age (mean, S.D), marital status (number, percent) and race/ethnicity (number, percent). For the fictitious quality improvement activity: is the QI intervention *associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* It is important to understand what is known about the people who smoked including gender (number, percent), age (mean, S.D), and race/ethnicity (number, percent). For the fictitious program evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program? Describe the demographic characteristics of the people with diabetes in the sample by age (mean, S.D), gender (number, percent) and race/ethnicity (number, percent). For the fictitious study of home visiting program outcomes of individuals receiving brief family home visiting services compared to those receiving long term visits, no characteristics the individuals within the sample were available. The sample description can provide the frequencies (number, percent) of individuals overall and in groups defined by the number of visits for each individual (1–2 visits group and 3 or more visits group) for three data sources (A, B, C).

14.2.3 Intervention

Consider options for analysis of the interventions as described in Chapters 5 and 6, to provide an understanding of the intervention associated with the outcomes. To describe the intervention:

- Use frequencies to describe intervention counts or percentages (number, percent) [15, 16]
- Use ratios to compare proportions of interventions in identified clusters or groups (percent) [15, 16]
- Use cross-tabs with Chi-Square (χ^2) to test the significance of difference in intervention counts (*p*-value) [15, 16]

For example, for the fictitious intervention effectiveness research project: are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?* Describe the average intervention count per mother (number, percent) and content (number, percent, ratio) overall and/or by problem or visit. For the fictitious quality improvement activity: is the QI intervention *associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* Describe the intervention content and the number of times each person who smoked cigarettes on admission to care who received the intervention. For the fictitious program evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program? Describe the program content and the average participation in program activities per person (number, percent). For the fictitious study of intervention patterns associated with outcomes of adult and adolescent mothers with and without the mental health problem, describe the sample for each of the intervention approaches (number, percent) as well as comparisons of outcomes by intervention approach (*p*-value).

14.2.4 Outcome

Consider options for analysis of the outcomes as described in Chapters 5 and 6. For outcomes analysis (P_{Time2} and P_{Time2}-P_{Time1}):

- Use standard descriptive statistics to describe the severity of the problem before before-and-after intervention (P_{Time2}, P_{Time1}) (Mean, S.D.) [15, 16]
- Use parametric test to evaluate the significance of differences between before-and-after measures (P_{Time2}-P_{Time1}) when the data meets the assumptions for each test (e.g. normality, homogeneity, independence) (*p-value*) [15, 16]
- Use a non-parametric test to evaluate the significance of differences between before-and-after measures (P_{Time2}-P_{Time1}) when the data does not meet the assumptions for an typical parametric test (*p-value*) [15, 16]
- Use a measure of effect size to examine magnitude of a difference (P_{Time2}-P_{Time1}) (e.g. Cohen's d) [17, 18]
- Use generalized estimating equations or other standard methods of assessing the difference between differences (P_{Time2}-P_{Time1} for group 1 compared to P_{Time2}-P_{Time1} for group 2) (*p-value*) [20]
- Use survival analysis and Kaplain Meier curves to assess time-to-outcome [15, 19, 21, 22]

For example, for the fictitious intervention effectiveness research project: are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?* Describe health equity (health status measure) at baseline and after discharge (mean, S.D.). Describe percentage of mothers who attained a benchmark

(P_{Time2}) and the significance of attaining a benchmark (e.g. generalized estimating equations). Report differences between baseline and final health equity (P_{Time2}-P_{Time1}) (e.g. paired samples *t*-test) (*p*-value). For the fictitious quality improvement activity: is the QI intervention *associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* Describe number of cigarettes smoked per day at baseline and after QI intervention (mean, S.D.). Describe benchmarks of people who smoked who quit or cut down to reach an established goal (number, percent) (P_{Time2}). Report differences between baseline and final measures (P_{Time2}-P_{Time1}) (e.g. cross tab, Chi-Square (χ^2)) (*p*-value). For the fictitious program evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program? Describe BMI at baseline and after program (mean, S.D.). Describe benchmarks of people with diabetes who decreased BMI or reached an established goal (number, percent) (P_{Time2}). Report differences between baseline and final BMI measures (P_{Time2}-P_{Time1}) (*p*-value). For the fictitious study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions, describe associations of knowledge, behavior, and status outcomes (P_{Time2} and P_{Time2}-P_{Time1}) with frailty and social determinants and interactions between frailty and social determinants (e.g. two way ANOVA) (*p*-values).

14.2.5 Relationships Among Variables

Consider options for analyzing relationships between variables that may reveal associations between PIO MM variables. As with EDA, planning these analyses may be facilitated using the matrix provided in Chapter 7 to arrange PIO MM variables in comparison with other PIO MM variables. Methods for analyzing relationships among multiple variables include:

- Use correlation to test significance of associations between variables (*p*-value) [15, 16]
- Use ANCOVA to account for a continuous variable in testing significance of mean differences (*p*-value) [15, 16]
- Use regression to account for numerous mean differences in modeling the proportion of an outcome that may be attributable to other factors and the significance of those factors (*p*-value) [15, 16]

For example, for the fictitious intervention effectiveness research project: are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?* Evaluate the proportion of the health equity outcome that may be attributable to the number or type of public health nursing interventions (e.g. hierarchical multiple regression model, accounting for demographics and baseline scores). For the fictitious quality improvement activity: is the QI intervention *associated with change in behavior of people who smoked cigarettes on admission to*

care after our QI PDSA cycle? Evaluate the association between number of cigarettes per day and benchmark attainment (e.g. Spearman's rho) (*p*-values). For the fictitious program evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program? Evaluate the association between extent of program participation and BMI reduction (e.g. regression techniques suitable for the variables) (*p*-values). For the fictitious study of intervention patterns associated with outcomes of adult and adolescent mothers with and without the mental health problem, evaluate the correlation between the number of problems and the number of interventions (e.g. Pearson's *r*) (*p*-value).

After completing Worksheet E, ensure that ethical permissions are in place (Chapter 8), and review the entire plan with stakeholders. Then implement the project: collect data or obtain data, preprocess, and conduct the analysis. After these steps are completed, move on to Chapter 15 to complete the results statements.

Reflection Questions

- How does the project purpose address the gap in knowledge? Relate to the hypothesis? To the research question?
- What is the value of defining variables that operationalize the PIO MM? How do the variables inform the design and analysis plan?
- What statistical methods are needed to describe benchmark and change outcomes for your project? Why?
- What statistical methods are needed to relate interventions to outcomes for your project? Why?

References

1. Centers for Disease Control and Prevention [US]. Framework for program evaluation in public health. MMWR Sept 17, 1999/48 (RR11); pp. 1–40
2. Institute for Healthcare Improvement. The science of improvement: how to improve. [Internet]. [cited 2017 May 12]. Available from: http://www.ihi.org/resources/pages/howtoimprove/scienceofimprovementhowtoimprove.aspx
3. Lee S, Kim E, Monsen KA (2015) Public health nurse perceptions of Omaha System data visualization. Int J Med Inform 84(10):826–834
4. Kirk A (2012) Data visualization: a successful design process. Packt Publishing Ltd, Mumbai
5. Gehlenborg N, Wong B (2012) Points of view: heat maps. Nat Methods 9(3):213
6. Aigner W, Miksch S, Schumann H, Tominski C (2011) Visualization of time-oriented data. Springer, Berlin
7. Byron L, Wattenberg M (2008) Stacked graphs–geometry & aesthetics. IEEE Trans Vis Comput Graph 14(6):1245–1252
8. Inselberg A (2009) Parallel coordinates: visual multidimensional geometry and its applications. Springer, New York, pp 1–554
9. Kerr MJ (2013). Mobile technology and GIS maps of sign/Symptoms data. In: First international conference on research methods for standardized terminologies, Apr, 2013, West St. Paul, MN

10. Monsen KA, Peterson JJ, Mathiason MA, Kim E, Votova B, Pieczkiewicz DS (2017) Discovering public health nurse–specific family home visiting intervention patterns using visualization techniques. West J Nurs Res 39(1):127–146
11. Harter JM, Wu X, Alabi OS, Phadke M, Pinto L, Dougherty D, Petersen H, Bass S, Taylor II RM (2012) Increasing the perceptual salience of relationships in parallel coordinate plots. In: InIS&T/SPIE Electronic Imaging, 2012 Jan 22. International Society for Optics and Photonics, pp 82940T–82940T
12. Few S Multivariate analysis using parallel coordinates. Perceptual Edge web site. [Internet]. [cited 2017 May 12]. Available from: http://www.perceptualedge.com/articles/b-eye/parallel_coordinates.pdf. Published Sept 12, 2006. Updated 2015
13. Radhakrishnan K, Monsen KA, Bae SH, Zhang W (2016) Visual analytics for pattern discovery in home care. Appl Clin Inform 7(3):711–730
14. Smith AK, Ayanian JZ, Covinsky KE, Landon BE, McCarthy EP, Wee CC, Steinman MA (2011) Conducting high-value secondary dataset analysis: an introductory guide and resources. J Gen Intern Med 26(8):920–929
15. Dunn OJ, Clark VA (2009) Basic statistics: a primer for the biomedical sciences. Wiley, Hoboken, NJ
16. McDonald JH (2009) Handbook of biological statistics. Sparky House Publishing, Baltimore, MD
17. Cohen J (1992) A power primer. Psychol Bull 112(1):155–159
18. Johnson KE, McMorris BJ, Raynor LA, Monsen KA (2013) What big size you have. Using effect sizes to determine the impact of public health nursing inter-ventions. Appl Clin Inform 4:434–444
19. Monsen KA, Farri O, McNaughton DB, Savik K (2011) Problem stabilization: a metric for problem improvement in home visiting clients. Appl Clin Inform 2:437–446
20. Monsen KA, Radosevich DM, Johnson SC, Farri O, Kerr MJ, Geppert JS (2012) Benchmark attainment by maternal and child health clients across public health nursing agencies. Public Health Nurs 29(1):11–18
21. Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. J Amer Statist Assoc 53:457–481
22. Freedman DA (2012) Survival analysis. Am Stat:110–119

Interpret the Results (Worksheet F) 15

15.1 Preparing to Complete Worksheet F

Completing the analysis generates results that must be interpreted in alignment with what is known about the population in context, as well as study design, variables, and analysis approach [1–4]. In this chapter, a systematic process for reviewing and stating (presenting) the results of the analyses is presented, together with the alignment of the results with the scientific literature. A statement regarding statistical significance of the findings should be accompanied by a statistical result (e.g. *p-value*) [5]. A statement of clinical significance may be accompanied by a statistical test as well (e.g. Cohen's *d*); however, it is most important for the statement of clinical significance to address important clinical issues related to the project.

Statements of the results and related presentation approaches are important aspects of reporting the results. Each aspect of the PIO MM will be addressed, with examples from fictitious studies described in Chapters 11–13. Unbiased reporting may be accomplished using the guidance in the paragraphs below, and such statements may be less likely to be interpreted incorrectly. (For example, misstatements of the result as causal or generalizable may go beyond the results and infer interpretation for other contexts and populations [1]):

- *Interpretation of the results within context.* The results of all intervention effectiveness research, quality improvement activities, and program evaluation projects based on PIO MM are bound by the context, as is the intent of the project. Therefore, it is critical to provide details regarding the sample including groups within the sample, and likewise to provide descriptions of the context including location and interventionists. Ways to organize, classify, interrelate, compare, and display information depend on the study design and the variables (continuous or categorical). These decisions are guided by the questions being asked, and by input from stakeholders and clinical experts [1].
- *Interpretation of results given the single-group before and after design.* In studies that are not randomized controlled trials, it is not possible to infer causation

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_15

due to lack of controls. Therefore it is essential to avoid words that imply causation such as "effect" and "impact" [1–4]. Rather, words that describe relationships among observed values such as "positively correlated" or "associated" are more accurate descriptions of the results.

- *Alternative explanations.* As described in Chapter 3, the inability to impose controls in most intervention effectiveness research, quality improvement activities, and program evaluation projects threatens the validity of the results [1–4]. Rigorous reporting of the project results must therefore include a discussion of alternative explanations.

15.2 Results Statements and Presentation

The results statements are presented below as described in Chapters 5 and 6, and in Worksheet F and are intended to be used as templates to formulate results statements from PIO MM projects. They may be customized for each project. Note that each statement uses simple language to state each result from the perspective of the type of statistical test that was used (descriptive or inferential). The examples below are typical for studies that use simple variables specific to PIO MM with a single-group before-and-after design and standard descriptive and inferential statistics. Modifications specific to other analyses may be needed. Note that words that imply causation such as 'impact' or 'effect' of intervention on outcome are not used in the prototype statements. Alternative explanations may be entered into Worksheet F within the appropriate align with literature sections.

15.2.1 Presenting the Results

When comparing descriptive and inferential statistics regarding groups within a sample, descriptive statistics for the whole sample and each of the groups may be reported in narrative (for simple descriptions) and/or a table (for multiple characteristics and/or groups). Examples of tables formatted in APA style for reporting demographics characteristics of a sample and results of various analyses are provided in each of the following sections (Tables 15.1–15.5). These examples are fictitious and provide formatting guidance without reference to a study. Between group differences shown by inferential tests should be noted for all results. Results presented in narrative format may be more general, with details shown in a table. As a general rule, results presented in a table should be summarized in narrative in statements and not repeated numerically.

15.2.2 Description of Sample Characteristics

Statements such as the following may be used: The characteristics of the sample were XX (XX%) (categorical characteristics), (M = XX.XX, SD = X.XX) (continuous characteristics) [5, 6]. If applicable, report the significance of differences in

Table 15.1 Example of a table showing the description of a sample overall and by group (data are fictitious)

Demographic characteristics by group								
	Total $N = 486$		Groups					
			Group 1 $n = 165$		Group 2 $n = 321$			
	Mean	(SD) or %	Mean	(SD) or %	Mean	(SD) or %	Statistic	*p* Value
Age							$t = 2.46$	<0.05
Mean	23.4	(7.4)	24.9	(8.1)	23.3	(6.9)		
Range	14–59		15–54	14–59				
Sex							$x^2 = 3.26$	n.s.
Female	474	97.5%	158	95.8%	316	98.4%		
Male	12	2.5%	7	4.2%	5	1.5%		

n.s. = not significant

sample characteristics between groups: There were differences in (characteristic) between groups in the sample (as below, depending on the test used) [5, 6]. See example of sample characteristics in Table 15.1.

For example, for the fictitious intervention effectiveness research project: Are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?*

- There were 2434 individuals in the sample with an average age of 21.2 (SD = 4.7).
- The sample was primarily white (67%). [The number of individuals in the "white" demographic group does not need to be reported because the total number of 2434 was given in the previous sentence.]

For the fictitious quality improvement activity: Is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI cycle?

- There were 366 individuals who received the QI intervention over the course of 1 month.
- Over half were female (57%).
- There were differences by gender in individuals who received 1–2 QI interventions and those who received three or more QI interventions.

For outcome evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?

- There were 189 participants in the BMI reduction program.
- Of these, 51% were male.

- There were 151 (80%) participants that attended at least half of program sessions.
- Of these, significantly more (65%) were male (χ^2 (2, N = 189) = 0.56, p = 0.006).

For the fictitious study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions:

- The sample of 1613 older adults was evenly divided between individuals 65–80 (51%) and 81 or above.
- There were more significantly females (75%) (χ^2 (2, N = 1613) = 0.64, p = 0.026) than males in the sample overall, and by age group. For individuals 65–80, 60% were female (χ^2 (2, N = 802) = 20.45, p = 0.023) and for individuals 81 and above 90% were female (χ^2 (2, N = 811) = 39.59, p < 0.001).

15.2.3 Description of Interventions

Statements such as the following may be used: There were XXX interventions, with an average of (M = XX.XX, SD = X.XX) per (person, visit, problem). Most interventions were for the (problem, group) (XX%) [5–6]. Significance of intervention differences overall and by group or problem may be reported as applicable following the description of interventions: There were differences in intervention (types or characteristics) by (group, problem) depending on the test used [5, 6].

For example, for the fictitious intervention effectiveness research project: are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?*

- There were 289,646 interventions in the sample with an average per person of 119 (SD = 43).
- Most interventions (66%) were for the Caretaking/parenting problem (Table 15.2).

For the fictitious quality improvement activity: is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI cycle?

- The QI Intervention was received 1–2 times more often (70% of individuals), than three or more times (30% of individuals).

For outcome evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?

- The program consisted of ten sessions that provided standardized content based on the program as described in the literature.

Table 15.2 Example of a table showing the description of interventions overall and by group (data are fictitious)

Intervention frequencies by problem for all cases and by group						
			Group			
	Total $N = 51,341$		Group 1 $n = 12,590$ for 165 cases		Group 2 $n = 38,751$	
	486	100	165	33.95	321	66.05
Problem	n	Percent of all interventions	n	Percent for this problem	n	Percent for this problem
All problems	51,341	100	12,590	24.52	38,751	75.48
Caretaking/parenting	28,113	54.76	8206	29.19	19,907	70.81
Antepartum/postpartum	10,761	20.96	2750	25.56	8011	74.44
Income	5153	10.04	798	15.49	4355	84.51
Mental health	1929	3.76	101	5.24	1828	94.76
Residence	1687	3.29	244	14.46	1443	85.54
Substance use	1348	2.63	200	14.84	1148	85.16
Family planning	958	1.87	181	18.89	777	81.11

For the fictitious study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions

- There were 141,944 interventions in the sample with an average per person of 88 (SD = 57).
- Most interventions (53%) were for the Treatments and Procedures category.

15.2.4 Description of Outcomes

Statements such as the following may be used: On average, final outcomes (P_{Time2}) for the entire sample were (M = XX.XX, SD = X.XX), an average change of (P_{Time2}—P_{Time1}) (M = XX.XX, SD = X.XX) from the baseline score (P_{Time1}) (M = XX.XX, SD = X.XX) [5, 6]. Significance of outcome (P_{Time2} compared to P_{Time1}) overall: There were differences in outcome after intervention overall depending on the test used (*p*-value), with effect size of X.X (Cohen's *d*). Significance of differences in outcome (final scores, change scores) by group, problem: There were differences in (benchmark attainment, improvement) by (group, problem) (as below, depending on the test used) [5, 6]. If there was a change but it was not significant, it may be noted that there was a change in the expected direction that was evaluated, however, the change was not significant.

For example, for the fictitious intervention effectiveness research project: are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?*

- On average, overall final health status scores for the entire sample were 4.22 (SD = 0.77); and average positive change of 0.61 (0.23).

For the fictitious quality improvement activity: is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?

- On average, the number of cigarettes smoked per day decreased 32% for individuals receiving the QI intervention.

For outcome evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?

- Participants in the BMI reduction program reduced BMI values from 32.3 (10.1) to 28.6 (8.8) on average.

For the study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions

- Overall, adults receiving home care interventions improved in Knowledge (Final $t(2, 1613) = 3.55$, SD = 0.77, $p < 0.001$, $d = 0.13$; Change $t(2, 1613) = 0.27$, SD = 0.11, $p < 0.001$, $d = 0.15$), Behavior (Final $t(2, 1613) = 3.79$, SD = 0.67, $p < 0.001$, $d = 0.25$; Change $t(2, 1613) = 0.33$, SD = 0.17, $p < 0.001$, $d = 0.31$), and Status (Final $t(2, 1613) = 4.23$, SD = 1.77, $p < 0.001$, $d = 0.55$; Change $t(2, 1613) = 0.77$, SD = 0.31, $p < 0.001$, $d = 0.61$) following home care interventions. Examples of similar outcomes tables are provided (Tables 15.3 and 15.4).

Table 15.3 Example of a table showing the description of three outcomes overall and by group (data are fictitious)

Outcomes by group								
	Total $N = 486$		Group				Statistic	p Value
			Group 1 $n = 165$		Group 1 $n = 321$			
	Mean	(SD) or %	Mean	(SD) or %	Mean	(SD) or %		
Knowledge ratings								
Baseline	2.87	(0.55)	3.06	(0.60)	2.77	(0.49)	5.50	<0.001
Final	3.46	(0.57)	3.70	(0.63)	3.34	(0.50)	6.48	<0.001
Change	0.59	(0.44)	0.65	(0.46)	0.56	(0.44)	2.00	<0.05
Behavior ratings								
Baseline	3.48	(0.63)	3.79	(0.68)	3.32	(0.54)	7.84	<0.001
Final	3.90	(0.68)	4.24	(0.66)	3.73	(0.62)	8.30	<0.001
Change	0.43	(0.46)	0.44	(0.49)	0.42	(0.44)	0.65	n.s
Status ratings								
Baseline	4.01	(0.72)	4.42	(0.62)	3.80	(0.67)	9.88	<0.001
Final	4.36	(0.67)	4.74	(0.46)	4.16	(0.68)	10.97	<0.001
Change	0.35	(0.53)	0.32	(0.49)	0.36	(0.54)	−0.89	n.s

Table 15.4 Example of a table showing the comparisons of two outcomes (final and change) overall and by group (data are fictitious)

Between-group comparisons of mean outcome ratings by intervention type and group						
	Mean outcome ratings					
	Group 1		Group 2			
Outcome	Intervention type A $n = 107$	Intervention type B $n = 58$	Intervention type A $n = 138$	Intervention type B $n = 183$	F	p Value
Final outcome 1 (P_{Time2})	3.78_{ab}	3.56_{c}	3.35_{a}	3.32_{bc}	18.610	<0.001
Final outcome 2 (P_{Time2})	4.31_{de}	4.10_{fg}	3.83_{df}	3.66_{eg}	26.556	<0.001
Outcome 1 change (P_{Time2}–P_{Time2})	0.67	0.59	0.55	0.57	1.833	n.s.
Outcome 2 change (P_{Time2}–P_{Time1})	0.44	0.44	0.45	0.39	0.525	n.s.

n.s. = not significant
Note: Means having the same subscripts are significantly different using the Bonferroni multiple comparisons method ($p < 0.05$)

15.2.5 Description of Benchmark Attainment

A statement such as the following may be used: Of the entire sample, XX% attained the desired benchmark of XX (state the definition of the benchmark number) [5, 6].

For example, for the intervention effectiveness research project: are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?*

- Of the entire sample, 81% reached the desired overall benchmark of four in status.

For the quality improvement activity: is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?

- Of the entire sample, 31% reduced or quite smoking during the PDSA cycle.

For the outcome evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?

- Of the entire sample, 56% attained the personal goal in BMI reduction set at the beginning of the program.

For the study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions:

- Of the entire sample, 94% attained the desired overall benchmark of three in Knowledge, 51% attained the desired overall benchmark of four in Behavior, and 88% attained the desired overall benchmark of four in status.

15.2.6 Correlations Between Interventions and Outcomes

Statements such as the following may be used: Intervention (totals, types) were positively associated with outcomes (overall and/or by problem) depending on the test used (*p*-value). For regression of independent variables on outcome: Accounting for individual characteristics (demographics and baseline assessments), the interventions explained XX% of the outcome (R^2 = X.XX, F(X, XXX) = XX.XX, p = X.XX) [5, 6].

For example, for the intervention effectiveness research project: are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?*

- As interventions increased, health status change increased ($r(2343) = 0.37$, $p < 0.001$)

For the quality improvement activity: is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?

- Cutting down or quitting was associated with the diagnosis of a respiratory condition and receiving more QI interventions ($F(4, 366) = 23.66$, $p = 0.04$)

For outcome evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?

- Male participants who attended at least half of program sessions had significantly greater BMI reduction compared to male participants who attended less than half of program sessions (χ^2 (2, N = 96) = 14.06, p = 0.012). An example of regression results is provided (Table 15.5).

Table 15.5 Effects of intervention type, group, and interaction on final outcome (P_{Time2})

	df	*F* Value	*p* Value	Adjusted R^2
Corrected model	4189	36.62	<0.001	0.430
Intercept		84.17	<0.001	
Baseline of outcome (P_{Time1})		111.44	<0.001	
Intervention type		6.05	<0.05	
Group		3.31	n.s.	
Interaction (Intervention type - group)		1.81	n.s.	

n.s. = not significant

15.3 Results Interpretation

First, interpret findings by attributing meaning to the results. Interpret each outcome result separately. Then synthesize findings to detect patterns by combining sources of information to reach a larger understanding of agreement, convergence, or complexity. Examine descriptive statistics and the significance of final outcomes (P_{Time2}) and change in outcome (P_{Time2}–P_{Time1}) as well as benchmark attainment (%) overall and across groups. Look for patterns that are revealed in effect size and/or significance. Observe all results from a high-level perspective and explain in a brief sentence. Next, relate results to the expected results: How did the intervention effectiveness research, quality improvement activities, or program evaluation results align with expected results, and compare to the previous work upon which this project was based?

For example, for intervention effectiveness research: was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?

- A summary of outcomes across variables and groups: Health disparities between groups of high risk mothers based on race/ethnicity significantly decreased after public health nursing interventions. Risk relates to health disparities. Eighty percent of high risk mothers with low MRI scores attained the desired benchmark, compared to 40% of high risk mothers with high MRI scores. Interpretation: This is consistent with the literature of public health nursing intervention effectiveness, risk, and health equity.

For quality improvement: was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?

- A summary of outcomes across variables and groups: Significantly more individuals received the smoking cessation intervention and quit smoking during the QI PDSA cycle than in the prior year. Interpretation: The results are consistent with the evidence-based QI intervention results demonstrated in previous research.

For program evaluation: Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?

- A summary of outcomes across variables and groups: The BMI of people with diabetes decreased after program participation, with significantly greater reductions for males compared to females. O*f females with diabetes who participated in the BMI reduction program, only 25% attained their goal BMI during the program*. Interpretation: This is a surprising finding compared to previous research.

For the study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions:

- A summary of outcomes across variables and groups For the entire sample and all groups, Knowledge, Behavior, and Status benchmark attainment and change outcomes were influenced by frailty and social determinants. Interpretation: This is a novel finding for the population of older adults that should be replicated using other datasets.

15.3.1 Theoretical Framework-Related Interpretation

Using the PIO MM will enable description of 'what is' within the project data related to a theoretical framework. Stating the results as they align with a theoretical framework may lend support to further use of that framework in the future, and will allow for statements related to the framework (e.g. findings align with the XX theory of behavior change).

For example, for intervention effectiveness research: was there a difference in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?

- The finding of increased health equity among high risk mothers after public health nursing interventions aligns with the Integrated Theory of Health Behavior Change [7]

For quality improvement: was there a difference in the behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?

- the finding of differences in the behavior of people who smoked cigarettes on admission to care before and after our QI PDSA cycle align with the Theory of Planned Behavior [8]

For program evaluation: Was there a difference in obesity among people with diabetes who participated in the BMI reduction program?

- The finding of decreased obesity among people with diabetes who participated in the BMI reduction program aligns with Self Determination Theory [9]

For the study of the variability of outcomes associated with frailty and social determinants of older adults receiving home care interventions:

- The finding that outcomes were influenced by frailty and social determinants aligns with the Social Cognitive Theory [10]

These statements may be followed by comments describing the concepts in the theory that relate to the PIO MM concepts as operationalized in the project. Such statements of alignment strengthen findings by showing the relationship of the findings to previous scholarly work. This is not to be confused with Theory testing research; rather, a theory's validity should be studied using theory testing methods.

15.3.2 Temporality-Related Interpretation

Using the PIO MM may aid in understanding temporal sequences depending on how time is modeled in the project. If using longitudinal methods to understand temporality, statements of time-to-outcome may be appropriate (e.g. time to outcome differed

by problem) [11]. The examples discussed here generally incorporate time as a function of the outcome variable, in which there are at least two points in time in which the problem is assessed (measured); and between which interventions occurred. Additional data points are needed to conduct time-dependent analyses such as survival analysis [11, 12].

15.3.3 Give Alternative Explanations for the Findings

In most projects that are not randomized controlled trials results must be interpreted with caution, due to the numerous threats to validity and biases inherent within observational data, particularly for the single group before and after study design [1–4]. For such projects, alternative explanations for the findings should always be discussed. More than one well-supported interpretation of a single finding may exist, in which case each should be described in relationship to the literature [1]. Be sure to share the findings and your interpretation with the stakeholders—whenever possible these stakeholders should be involved in documenting data or providing health care to the population of interest. Doing so will provide a powerful litmus test for your interpretations, and allow you to report clinically relevant findings within their natural context. This increases your ability to have confidence in your conclusions [1]. It is important to validate these results through replication with other data sets [1–4].

For example, for the intervention effectiveness research project: are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?*

- When using existing data, there are many alternative explanations beyond what can be known from the dataset. For example, maturation is a threat whenever interventions take place over time.

For the quality improvement activity: is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?

- The motivation for the change in behavior is not known, and only short term findings can be reported so the duration of the change is not known.

For the program evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?

- There may be other reasons besides program participation for the BMI change, including increased activity due to the season of the year, or a seasonal change in diet.

Review the results statements and consider the meaning conveyed by each result as well as an overall message that can be known from the data. Remember to avoid causal statements. Explain the results to others in stories and pictures. Upon completing Worksheet F, the project is finished, and the dissemination process should begin.

Reflection Questions

- Based on the results how do the outcomes relate to the interventions? Why is this important for the population of interest?
- How do the results align with what is known about effectiveness of the intervention?
- What new patterns may be discovered through examining results across groups and variables? How would these patterns be validated?
- What does it mean to go beyond the data when interpreting results? What language should be used and what language should be avoided?

References

1. Centers for Disease Control and Prevention [US] (1999) Framework for program evaluation in public health. MMWR 48(RR11):1–40
2. Silverman B, Mai C, Boulet S, O'Leary L (2009) Logic models for planning and evaluation. The division of birth defects and developmental disabilities (DBDDD). Centers for Disease Control and Prevention (CDC), Atlanta, GA
3. Chapter 3 Single Group Before-and-After Design [Internet]. [cited 2017 May 12]. Available from: http://158.132.155.107/posh97/private/research/evaluation/Chp_3.pdf
4. Kernler D. Section 1.2: Observational Studies versus Designed Experiments.[Internet]. [cited 2017 May 12]. Available from: https://faculty.elgin.edu/dkernler/statistics/ch01/1-2.html
5. Dunn OJ, Clark VA (2009) Basic statistics: a primer for the biomedical sciences. Wiley, Hoboken, NJ
6. McDonald JH (2009) Handbook of biological statistics. Sparky House Publishing, Baltimore, MD
7. Ryan P (2009) Integrated theory of health behavior change: background and intervention development. Clinical Nurse Specialist (CNS) 23(3):161
8. Ajzen I (2011) Theory of planned behavior. Handb Theor Soc Psychol Vol One 1(2011):438
9. Deci EL, Ryan RM (2011) Self-determination theory. Handbook of Theories of Social Psychology 1:416–433
10. Bandura A (2001) Social cognitive theory: an agentic perspective. Annu Rev Psychol 52(1): 1–26
11. Freedman DA (2012) Survival analysis. Am Stat:110–119
12. Monsen KA, Farri O, McNaughton DB, Savik K (2011) Problem stabilization: a metric for problem improvement in home visiting clients. Appl Clin Inf 2:437–446

Disseminate the Findings

16

16.1 Why Dissemination Matters

Dissemination is the process of communicating unbiased findings to relevant audiences in a timely manner [1]. The goal of intervention effectiveness research, quality improvement activities, and program evaluation is to inform decision-making and guide appropriate action. Thus, findings must be disseminated appropriately in order to be implemented in practice, inform policy, and further advance science [1].

Expectations of stakeholders may differ regarding preferred dissemination methods, which may consist of a formal evaluation report, a scientific publication, or a presentation to administrators and staff. The dissemination strategy that is optimal for each project should be planned in advance in order to meet the needs of all stakeholders in the practice and scientific communities. It is important to consider timing, style, tone, message source, vehicle, and format of information in presentations, reports, and publications [1].

The completed Worksheets A–F will generally support development of most dissemination materials. This chapter will focus on development of a manuscript for publication in the scientific literature based on the project results.

The process of scientific publishing may seem daunting; and at the same time, it can be rewarding. You may be familiar with the adage "If you don't document it, it wasn't done"—the mantra of legal advisors to the nursing profession over many years. The corollary adage—"If you don't publish it, it wasn't done"—suggests that dissemination is an essential part of advancing scientific discovery. Publishing project results in the scientific literature means that the completed work builds evidence of intervention effectiveness, quality improvement, and program evaluation. Doing so enables others to identify and build on these results. Various other forms of dissemination such as conference presentations and academic theses may be less frequently found in literature searches. Therefore, it is essential that the findings of completed studies are published.

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_16

16.2 Getting the Most Benefit from This Chapter

As you begin this chapter, take a moment to reflect on everything that you have experienced throughout the project and what the results mean to you. Then think out loud about the whole project, describing it in simple language. Do this with co-authors, and also with others such as friends, colleagues, clinical experts, and mentors. Keep notes as you gain insights about the meaning of your findings. This process helps you synthesize and refine your interpretation and will lead to personal growth in understanding and result in a better story.

In this approach to dissemination, we describe sections of a manuscript that align with usual and customary publishing requirements of intervention effectiveness, quality improvement, and program evaluation studies, irrespective of journal or publication guidelines. The reason for our approach is twofold. First, there are many journals with variable author guidelines. Second, following a guideline is a rote process that does not support synthesis of the entire project into a meaningful story that will be of value to scholars and clinicians. We intend to make visible the thought processes that underlie manuscript development for development of a solid first draft. As manuscripts often have numerous revisions before they are submitted to a journal, formatting per author or publication guidelines may be completed after the manuscript has been drafted and is in a stable form.

Many times authors include co-authors who are team members or others such as advisors who have contributed substantively during the project. It is generally a good idea to invite co-authors and agree on authorship duties and order prior to writing the manuscript. Occasionally a co-author who was not involved during the project may assist with the publication by assisting with writing or statistical review. In some cases, a student may need to be sole author of a publication. In this case, it is considered appropriate to mention mentors in an acknowledgments section. Funding sources should always be acknowledged. The first author is considered responsible for the entire publication and process, and will do most of the actual writing.

The evaluation of intervention patterns associated with outcomes of adult and adolescent mothers with and without the mental health problem that was described in Chapter 13 is presented in this chapter in order to illuminate the iterative process that results in rapid, successful dissemination. Please review Chapter 13 Worksheets as a starting point for this chapter.

16.3 Iterative Interpretation and Explication of the Overall Story

To decide what to mention in a manuscript, first understand the important story that is the basis of all good manuscripts. Think about the critical message that should be shared based on your project. Take a moment to write out or explain verbally what was important about it in simple language. Start with drafting the story of your project that needs to be heard and known in order to advance science, practice, and policy. This may mean that a manuscript reports only a portion of the entire project.

Consider examples from this book:

- For the intervention effectiveness research project: *Are public health nursing interventions associated with changes in health equity* (*reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?* The story may be that improvement in health equity is associated with public health nursing interventions, especially for a particular group within the population.
- For the quality improvement activity: *Is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?* The story may be that the QI intervention was associated with higher intention to quit and fewer cigarettes per day after at least three encounters in which advice was provided.
- For the program evaluation: *Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?* The story may be that the program participants reduced BMI over the course of program participation.

For the evaluation of intervention patterns associated with outcomes of adult and adolescent mothers with and without the mental health problem, you might say:

PHN interventions make a difference for high risk mothers, but little is known about which interventions work best for adolescents and adults, particularly looking at differences for mothers with and without the mental health problem. We used a large dataset of problems, interventions, and outcomes to explore what interventions seem to make the most difference for adolescent and adult mothers with and without the mental health problem. There was differential improvement and outcome attainment after intervention. Overall, clients improved following PHN intervention, with some variability in outcome looking at the sample by intervention group.

As you describe your insights and answer questions about your project, you will refine your understanding and update your story. This process will help you solidify your understanding of the perspective and message that you will share as you develop abstracts, presentations, and manuscripts.

For example, your project may have several aims and multiple important findings. It is easy to get caught up in the details of the findings and lose sight of an overall message. For example, for the intervention effectiveness research project: A*re public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?*

- As interventions increased, health status change increased.
- Mothers with lower health status received more interventions.
- Health equity was improved, however disparities persisted.

For the quality improvement activity: *Is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?*

- Few people quit or cut down after only one or two encounters.
- Males quit or cut down more than females.
- Cutting down or quitting was associated with the diagnosis of a respiratory condition.

For the program evaluation: *Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?*

- Participants with higher BMI were more successful compared to those with slightly elevated BMI.
- Participants who attended less than 50% of program sessions did not change.
- Males were less likely to reduce BMI than females.

For the evaluation of intervention patterns associated with outcomes of adult and adolescent mothers with and without the mental health problems, the story may be that:

- The number of problems was positively correlated with the number of interventions: Public health nurses tailor interventions to meet unique needs of individuals.
- Overall, adolescents were more likely to improve than adults: Age matters.
- Adults with the mental health problem were less likely to reach the desired benchmark compared to adults without the mental health problem: For adults, having the mental health problem matters.

After conversation with colleagues and clinical experts, you may discover that there are other important implications and interpretations, alternative explanations, and limitations. For example, for the intervention effectiveness research project: *Are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?*

- As interventions increased, health status change increased: Dose of interventions matters.
- Mothers with lower health status received more interventions: Public health nurses tailor interventions.
- Health equity was improved, however disparities persisted: Health disparities are a challenging issue in health care and society.

For the quality improvement activity: *Is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?*

- Few people quit or cut down after only one or two encounters: Keep asking about smoking even when no changes are occurring.
- Males quit or cut down more than females: Gender matters.
- Cutting down or quitting was associated with the diagnosis of a respiratory condition: Health condition matters.

For the program evaluation: *Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?*

- Participants with higher BMI were more successful compared to those with slightly elevated BMI: Degree of elevated BMI matters.
- Participants who attended less than 50% of program sessions did not change: Program participation matters.
- Males were less likely to reduce BMI than females: Gender matters.

For the evaluation of intervention patterns associated with outcomes of adult and adolescent mothers with and without the mental health problem:

- The four groups may represent intervention tailoring as the sample differed by intervention group in the number of problems and interventions per individual. As this is an observational study it is not possible to differentiate between intervention groups for causation of outcome, rather we describe what happened and propose possible interpretations.
- The four groups may represent different programs that are offered through the Public Health department, and these programs may serve group of individuals with differing risk levels, which would likely relate to final outcome attainment.
- The best outcomes were found in the intervention group with most teaching, guidance, and counseling interventions. Note that this does not necessarily mean that a higher proportion of Teaching, guidance, and counseling interventions will improve outcomes for all; rather, it may describe that individuals in the intervention group received tailored interventions that were optimal for the group. As with all observational data, unless there is a prescribed intervention protocol, it is necessary to assume that health care interventions are tailored to meet individual needs. Thus interventions and outcomes must be evaluated as a whole in relationship to individual characteristics and context, compared to all others as shown in the PIO MM [2, 3].

Keep a list of your main messages, and revise it frequently based on feedback you receive that will help you refine the messages. After several conversations, you will be at a point where you decide you have clarity about the story and you will be ready to begin your formal writing. Be sure to double check Worksheet F and review all worksheets to ensure that you have not missed important information. Keep in mind that there may be findings in Worksheet F that are not relevant or important for this story and manuscript, but may be useful for another.

Note that in every project, there may be more than one important message, so more than one manuscript may be needed to describe the findings. For example, there may be one aim that describes a quantitative outcome and another aim that describes a qualitative outcome. Both aims may have important findings that could be the basis of an entire manuscript. If this is the case, then draft the story for each of the manuscripts and work on them side by side so that the content can be complimentary and minimize duplication.

16.4 Drafting the Abstract: Summarize the Story in Brief

Using the refined list of main messages derived from conversations about the findings and Worksheets A–F, as described above, draft a brief abstract that tells your story. It is often helpful to incorporate headings to help structure the abstract. Later you can add numbers and significance of findings, and expand word numbers depending on the required format of your conference or journal.

For the fictitious evaluation of intervention patterns associated with outcomes of adult and adolescent mothers with and without the mental health problem, you might say:

Background: Public health nurse interventions improve outcomes for high risk mothers, but little is known about which interventions work best for particular populations. The purpose of this project was to examine intervention effectiveness for adolescents and adults, particularly looking at differences for mothers with and without the mental health problem.

Method: We used a large dataset of problems, interventions, and outcomes to explore differential associations between intervention groups and outcomes for adolescent and adult mothers with and without the mental health problem.

Results: Aim 1 found that differed by proportion of Teaching, guidance and counseling, Case management, and Surveillance. Aim 2 found that individuals improved following PHN intervention, with variability between intervention groups. Knowledge outcomes were significantly better for intervention group 3. Individuals in intervention group 3 also had the fewest problems and fewest mental health interventions, but had the second highest number of interventions. Aim 3 found outcome variability was primarily related to age. Between group analysis showed significant interactions for only the Knowledge outcomes. Adolescents in intervention group 1 had greater improvement in Knowledge than adults across intervention groups. Adults with the mental health problem had lower Knowledge change compared to adults without the mental health problem in intervention groups 1 and 2.

Conclusions: This project demonstrated that public health nurses tailored interventions to adult and adolescent mothers with and without the mental health problem. The findings should be validated in other datasets and in prospective research; especially to examine the role of Case management interventions for all groups.

16.5 Develop and Display Results

First, think about the main take-home messages. How can these messages be conveyed optimally in words, tables, or figures? If the problem improved after intervention, a figure depicting baseline and final values may tell the story. If benchmarks were attained, a table that displays the benchmark analysis may tell the story. Such tables and figures will be the centerpiece of the results section, framed by descriptive data regarding individual characteristics and intervention descriptions. Report the findings as in the statements in Worksheet F. Give main findings in the narrative and refer to tables and figures for more detailed findings. It is not necessary to report all findings if they do not add to the overall message. However, contradictory findings should not be hidden.

Examples from the abstract above could be expanded or depicted by aim as follows:

Aim 1 Results. Four groups differed by proportion of Teaching, guidance and counseling, Case management, and Surveillance (See Fig. 16.1).

Aim 2 Results. On average individuals improved following PHN intervention, with variability between groups in each IA. Knowledge outcomes were significantly better in IA3 (See Fig. 16.2).

Individuals in intervention group 3 also had the fewest problems and fewest mental health interventions, but had the second highest number of interventions (See Table 16.1).

Aim 3 Results. Outcome variability was primarily related to age. Between group analysis of all 16 groups of adolescent and adult mothers with and without the mental health problem in each intervention group by all other groups (N = 136 paired comparisons) showed significant interactions for only the Knowledge outcomes. The heat map/tables in Fig. 16.3 compares the total number of paired contrasts (N = 23, 16.9%) that were significant by intervention group counting from the directions of the x and y axes. Most significant contrasts were noted between adolescents

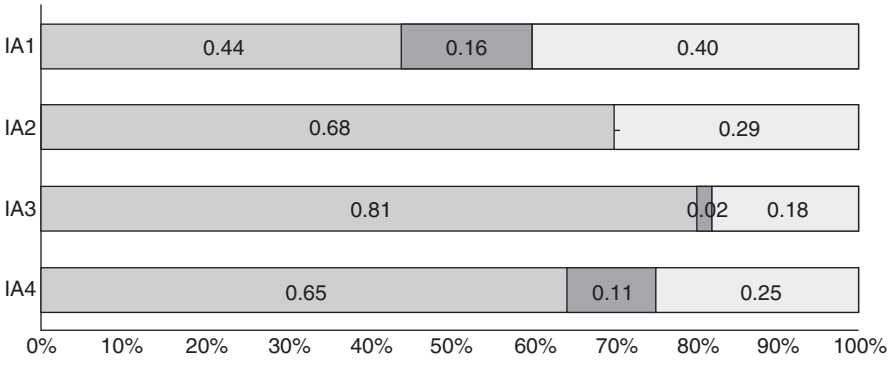

Fig. 16.1 Intervention groups by proportion of Teaching, guidance, and counseling (*left*), Case management (*center*), and Surveillance (*right*)

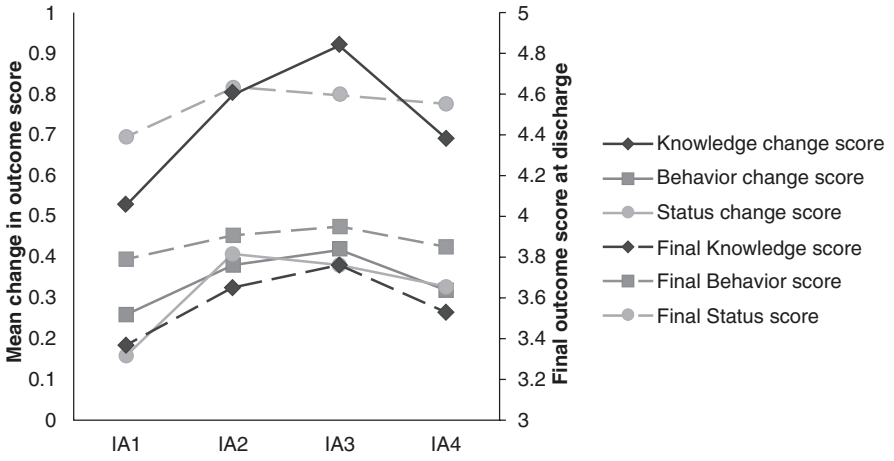

Fig. 16.2 Final and change Knowledge, Behavior, and Status scores by Group

Table 16.1 Numbers of problems, interventions, and mental health interventions by Intervention group

Between-group comparisons of numbers of problems, interventions, and mental health interventions by intervention group						
Characteristic	Intervention group 1 $n = 107$	Intervention group 2 $n = 58$	Intervention group 3 $n = 138$	Intervention group 4 $n = 183$	F	p Value
Number of problems	4.58	4.05	3.22	4.81	16.521	<0.001
Number of interventions	85.2	99.6	107.8	151.0	22.356	<0.001
Number of mental health interventions	5.18	2.03	0.20	3.82	31.773	<0.001

Significant contrasts (23 of 136 = 16.9%)

Counts for X axis	Total	Group 1	Group 2	Group 3	Group 4
adolescent-adult contrasts by MH/MH or no MH/no MH	9	1	2		6
adolescent-adult contrasts by MH/no MH	9	1	2	1	5
adult-adult contrasts by MH/no MH	2		1	1	
adult-adult contrasts by MH/MH or no MH/no MH	3		1	2	

Counts for Y axis	Total	Group 1	Group 2	Group 3	Group 4
adolescent-adult contrasts by MH/MH or no MH/no MH	9	6	2		1
adolescent-adult contrasts by MH/no MH	9	6	2		1
adult-adult contrasts by MH/no MH	3	2	1		
adult-adult contrasts by MH/MH or no MH/no MH	2	2			

Fig. 16.3 Knowledge differences by group among adolescents and adults with and without the mental health problem (1 = significant difference between groups at $p < 0.05$, 0 = no significant difference between groups). *MH* mental health

and adults (top rows of each table) in Groups 1 (lower table) and 4 (upper table). The fewest significant contrasts were found in Group 3.

These example statements show the importance of drafting results to describe the most important messages of the project and then developing tables and figures to display the results in a manner that makes the overall message obvious and clear. Ask several colleagues and clinical experts to read and interpret the draft results section to ensure that the message is correct and clear; and continuously revise as needed. Review the final results section verbally with someone who has not heard about it before, and clarify if need. When others understand your main message and you are confident that you have included the relevant findings, you have finalized the results section.

The results section is the centerpiece of the manuscript, and other sections must reflect and align with it. Additional information about findings that were not relevant to the story should not be included or discussed, except to mention that they are reported elsewhere if there is a reason to do so.

16.6 Adding Meaningful Interpretation to the Results

The discussion section explains the results in relationship to the literature. It is an extension of the story that allows for you to underscore the importance of the messages you chose to describe in the results. The next step is to build the discussion section based on the results section by adding insights that demonstrate how the project adds to the knowledge base for intervention effectiveness, quality improvement, or program evaluation. The basic structure of the discussion section starts with a summary paragraph, followed by a paragraph for each major finding, and then paragraphs that describe implications, limitations, and any needs for further research.

First, write a summary paragraph that describes the project purpose and findings and concludes with a statement of future research based on this project. This paragraph should consist of only a few sentences as you will go into detail in later paragraphs. It may be similar to the abstract, starting with the purpose of the project restated in the past tense.

For example, for the intervention effectiveness research project: *Are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?*

In this project we examined health equity by comparing health status indicators for low and high risk mothers who received public health nursing interventions. Overall, health equity increased. As interventions increased, health status change increased. Mothers with lower health status received more interventions. Health disparities between low and high risk mothers persisted, and are a challenging issue in health care and society. Further research is needed to validate these findings in other datasets.

For the quality improvement activity: *Is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?*

In this PDSA QI intervention cycle for advice to people who smoked cigarettes during each encounter, the average number of cigarettes per day was reduced after three encounters. Few people quit or cut down after only one or two encounters. Males quit or cut down more than females. Cutting down or quitting was associated with the diagnosis of a respiratory condition. Additional PDSA cycles are needed to incorporate gender- and condition-specific evidence-based interventions and to extend the QI intervention to other units and settings.

For the program evaluation: *Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?*

In this program evaluation, people with diabetes with elevated BMI who participated in at least 80% of the program were successful in reducing BMI. Participants with higher BMI were more successful compared to those with slightly elevated BMI, while participants who attended less than 50% of program sessions did not change. Males were less likely to reduce BMI than females. Program outcomes aligned with previous program effectiveness reports.

For the evaluation of intervention patterns associated with outcomes of adult and adolescent mothers with and without the mental health problem:

In this project we discovered four data-driven groups and examined outcomes relative to the interventions for adolescent and adult mothers with and without the mental health problem. Findings showed that public health nurses tailored interventions to adult and adolescent mothers with and without the mental health problem. On average individuals improved following intervention, with variability between groups. Outcome variability for all intervention groups was primarily related to age. Between group analysis showed significant interactions for only the Knowledge outcomes. The intervention groups should be validated in other datasets and in prospective research.

The next paragraphs describe each of the findings, their alignment with the literature, and the clinical importance of the findings. Prepare one such paragraph for each of the findings. Recommendations for further research, for practice, and or for policy may be offered based on the findings, either within the paragraphs for a specific finding or as a separate paragraph. Keep in mind that recommending continuing, expanding, modifying, or stopping an intervention or program is different from demonstrating intervention or program outcomes in a particular population or setting [1].

For example, for the intervention effectiveness research project: *Are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?*

The finding that as interventions increased, health status change increased aligns with the literature on the need for sufficient intervention dose to affect change in health outcomes [4]. This is important clinically because interventions are time consuming and costly, and programs need justification to provide sufficient interventions.

The finding that mothers with lower health status received more interventions aligns with the literature on intervention tailoring [5–8]. These two findings support public health nursing policies to provide more interventions when needed for those who need them.

The finding that health equity was improved but disparities persisted aligns with the literature on persistence of health disparities in health care and society [8]. That public health nursing interventions are associated with improved health equity is remarkable and suggest that health equity outcomes should be evaluated and documented when planning intervention effectiveness research.

For the quality improvement activity: *Is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?*

The finding that few people quit or cut down after only one or two encounters aligns with the literature on the need for sufficient intervention dose to affect change smoking behavior [9]. This is important clinically because interventions are time consuming and costly, and programs need justification to provide sufficient interventions.

The finding that males quit or cut down more than females aligns with the literature on the need to tailor interventions by gender [9]. The next PDSA cycle should incorporate evidence-based gender-specific interventions.

The finding that cutting down or quitting was associated with the diagnosis of a respiratory condition aligns with the literature on motivation for behavior change among people who smoke [9]. That the QI intervention was associated with changes in smoking behavior was remarkable and suggests that the QI intervention should be evaluated in PDSA cycles throughout the organization.

For the program evaluation: *Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?*

The finding that participants who attended less than 50% of program sessions did not change aligns with the literature on the need for sufficient intervention dose to affect BMI change [10]. It is important to promote full participation in the

program in order to provide sufficient interventions to achieve program outcomes.

The finding that males were less likely to reduce BMI than females aligns with the literature on the need to tailor interventions by gender [10]. Future programs should incorporate evidence-based gender-specific BMI-reduction interventions.

The finding that participants with higher BMI were more successful compared to those with slightly elevated BMI aligns with the literature on successful BMI reduction [10]. That the program was associated with changes in smoking behavior reinforces program use to address the critical challenges of reducing BMI among people with diabetes. This population health strategy can be expected to improve overall health of people with diabetes who participate fully in the program.

For the evaluation of intervention patterns associated with outcomes of adult and adolescent mothers with and without the mental health problem:

The finding that on average individuals improved following PHN intervention, with variability between intervention groups, particularly for Knowledge outcomes aligns with literature on public health nursing outcomes [5–8], and expands the knowledge base regarding interventions for adolescent and adult mothers with and without the mental health problem.

The finding that individuals in Group 3 had the highest final Knowledge scores, and also had the fewest problems and fewest mental health interventions is consistent with previous research regarding the challenges faced by mothers with mental health [6, 7, 11]. The finding that these individuals had the second highest number of interventions is surprising, and may reflect the need for more interventions to achieve higher outcomes, even for families with relatively fewer problems [8], or the assignment to a program that emphasizes frequent visits and numerous interventions [4].

The finding that outcome variability was primarily related to age aligns with previous literature that emphasizes the importance of early intervention with at-risk mothers [4] and supports the policies that prioritize this population.

The finding that adults with the mental health problem had lower Knowledge outcomes across intervention groups underscores the challenges faced by this population and aligns with the well-documented need for additional mental health resources in community settings to address this critical problem [12–14].

Additional paragraphs discuss implications for practice, research, and policy; and expand on other important aspects of the unique story. For example, for the intervention effectiveness research project: *Are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?*

Implications for practice for this project include reinforcing intervention tailoring to address the diverse needs of high risk mothers, while providing the information

and support needed by low risk mothers. In addition, public health nurses and managers alike should participate in data-interpretation sessions and training to learn from the findings and be aware of the value to interventions associated with increased health equity.

For the quality improvement activity: *Is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?*

Implications for practice are of primary importance in the Plan-Do-Study-Act cycle, incorporating the project (Study) findings into practice (Action). The consistent practice of advising people who smoke cigarettes to quit should be incorporated within each encounter. Further PDSA cycles should extend this intervention to other units. Training should accompany the intervention roll out to ensure the intervention is delivered correctly [9].

For the program evaluation: *Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?*

Implications for practice in population health programs are consistent with the literature on the importance of establishing and maintaining relationships with participants to encourage consistent attendance throughout the program [10]. Furthermore, the program outcomes were best for participants who experienced the full intervention as delivered by the interventionists who had received the training. Therefore it is important to provide training for program interventionists to ensure the intervention is delivered correctly [10].

Implications for Research may be described as follows in a new paragraph, or added to existing paragraphs with similar content. For example, for the intervention effectiveness research project: A*re public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) among high risk mothers after public health nursing interventions?*

Further research is needed to test intervention effectiveness among other populations served by public health nurses, and to understand reasons for differential intervention effectiveness between high risk mothers and low risk mothers.

For the quality improvement activity: *Is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?*

Further research is needed to discover reasons for gender- and condition-specific differences in smoking behaviors for people who receive healthcare advice to quit smoking. Such research may inform future QI interventions in the future.

For the program evaluation: *Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?*

Short-term program effectiveness was demonstrated in this project. Further research is needed to examine long term effects of the program with large samples of program participants vs. non-participants. Participants should provide feedback about the program and its impact over time, as well as aid in interpreting findings. Such research may provide insight into potential program changes.

For the evaluation of intervention patterns associated with outcomes of adult and adolescent mothers with and without the Mental health problem:

The round-up round-down method was successful in creating four intervention groups with different characteristics and enabling knowledge discovery on differential intervention delivery and associated outcomes [15]. This new method should be tested with other data sets and populations in order to validate these results and further refine the method.

Implications for policy may be described as follows in a new paragraph, or added to existing paragraphs with similar content. For example, for the fictitious intervention effectiveness research project: Are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?*

- This project supports national policy to provide home visits to high risk mothers to reduce health disparities and promote health equity [8]. Efforts to collect data for the evaluation of health disparities, including the social and behavioral determinants of health are being advanced through the Institute of Medicine [16]. However, quality documentation of these concepts is limited, and improvements in health policy are needed to advance the capacity to support similar analyses nationally.

For the quality improvement activity: is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?

- This project was successful in improving practice related to screening of people who smoked cigarettes on admission to care because policies were in place supporting the evidence-based QI initiative. These policies should be extended to support further spread of the QI intervention to other units and settings within the health system.

For the program evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?

- The successful outcomes of the BMI reduction program increased confidence in the effectiveness of the intervention and local policy makers were encouraged to support continuing the program for the next year, with particular emphasis on recruiting more males into the program.

For the evaluation of intervention patterns associated with outcomes of adult and adolescent mothers with and without the mental health problem:

- This project supports national policy to provide PHN home visits to high risk mothers of all ages with and without the mental health problem, as Knowledge, Behavior and Status outcomes improved for all groups, and behavior and status did not differ across all groups regardless of IA. There is a need to further extend the ability to conduct such studies through the reuse of PHN home visiting data [17]. This ability depends on quality documentation in EHRs that enable data aggregation and reporting. Local and state policies supporting such documentation are critical to advancing this agenda [18].

It is essential to discuss limitations of the single group before and after design and variations of the design, especially when observational data are used (as described in Chapters 2 and 3). In particular it is important to explain the fact that all large datasets are biased according to the observer perspective, and the lack of a control group makes it impossible to claim causation [19–21]. Replication of studies increases the confidence in findings.

16.7 Limitations

While it is important to recognize the limitations of a project, it is also critical to frame limitations in a way that does not negate the findings and value of the project. Thus, a limitation may be noted and followed immediately by a description of how that limitation was addressed in the project; and a statement regarding future research that would examine the possible bias introduced by the limitation. This paragraph may be revised to incorporate literature that applies to any of the examples:

- The limitations of this project are typical among retrospective comparative studies re-using EHR data. All such datasets are inherently biased by the documentation process [19, 20]. This project used a comparison approach to describe relative differences between four groups, and as such provides contrasting findings to improve interpretation. The alignment of the results with literature describing studies using other datasets [5–8], known community needs [12–14], and randomized trials [4] increases confidence in the findings.

Additional statements or paragraphs may be added to describe any recommendations that are justified by the findings. Recommendations for further research may

be made in any paragraph in which additional research is needed or desirable to validate the findings of the project. Recommendations for changes in practice should be supported by literature regarding the evidence-based intervention, and should be accompanied by a statement regarding further evaluation in the practice context. Such recommendations should be vetted by stakeholders [1].

16.8 The Methods Section

After completing drafts of the Results and Discussion sections, use the relevant aspects of Worksheet E to describe project methods in sufficient detail for replication of the project. The methods section begins by describing the study design and the ethical approvals. For example:

This retrospective, comparative study was approved by the University Institutional Review Board. We employed new data management techniques and used an existing dataset of de-identified interventions and outcomes.

Describing the methods in sufficient detail for replicating the study means that each step of the method could be repeated by an independent investigator with a similar dataset, using the methods description within the section together with references for that method. Well known methods may be summarized briefly (e.g. standard descriptive and inferential statistics). Instruments are described with their psychometric properties. For example:

Instrument: The Omaha System. The Omaha System [22] is a standardized terminology and ontology that describes problem concepts (Problem Classification Scheme), interventions (Intervention Scheme), and outcomes (Problem Rating Scale for Outcomes). All interventions and outcomes relate to a central concept called a Problem. There are 42 Problems in the Omaha System. Each Omaha System intervention is directed toward a single Problem, and consists of a Category term (Teaching, guidance, and counseling, Treatments and procedures, Case Management, and Surveillance); a defined Target term (n = 75); and includes a customizable field for further care description. Outcomes may be measured through use of the Problem Rating Scale for Outcomes, which consists of three valid and reliable Likert-type ordinal scales for the dimensions of Knowledge (1 = no knowledge, 5 = superior knowledge) Behavior (1 = not appropriate, 5 = consistently appropriate) and Status (1 = extreme signs/symptoms, 5 = no signs/symptoms). Validity and reliability of the Omaha System were established during its development [22].

Intervention groups: Novel intervention management methods were used to examine patterns in the intervention data [15]. First, a data driven approach was used to classify intervention data for each individual based on calculated percentages of interventions compared to the total number received, using a round-up, round-down algorithm. The resulting four patterns were then used to create four groups within the sample.

Data analysis: Bar and line graphs were created in Excel to visualize characteristics of the sample within the intervention groups. Inferential statistics were used to

test the significance of patterns relating individual characteristics, problems, categories, and outcomes associated with intervention groups using one-way ANOVA tests and two-way ANOVA tests with interactions in SAS v9.4 [23].

After drafting the methods section, review the results section again to make sure you have included the necessary descriptions and references. Then review the discussion section to check for alignment of methods, results and discussion. If you find inconsistencies among these sections, revise to make sure you are telling your story clearly, succinctly, and completely.

16.9 The Purpose Statement

After completing Methods, Background, and Discussion sections, review the purpose statement from Worksheet E for alignment with the manuscript story and results. Review the project purpose statement from Worksheet E, and revise as needed for the manuscript. For example:

The purpose of this project was to examine intervention effectiveness for adolescents and adults, particularly looking at differences for mothers with and without the mental health problem.

This language may be slightly modified to better explain the story of associations among the four intervention groups for adolescent mothers with and without the mental health problem:

The purpose of this project was to examine associations between interventions and outcomes of high risk adolescent and adult mothers with and without the mental health problem.

16.10 Background to Set the Stage for the Purpose

When the Purpose, Methods, Results, and Discussion sections align, the Background section may be completed based on the literature review. The Background section should usually contain substantive information related to the literature that is referenced in the Discussion section. As a general rule, most or all references should be introduced for the first time in the background and methods section.

The background section begins with a description of the importance of the problem or gap identified in Worksheet B, using the literature review. Most sentences in the background section should be supported by one or more references. Make a list of the important aspects of the problem and address each one with a paragraph supported by the literature. If the instrument is an important part of the project, it may be introduced in the background section with a comment that more information about the instrument is provided in the methods section.

For example, your list of paragraphs may include for the intervention effectiveness research project: Are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?* Adding a section regarding the reuse of EHR data in research (Topic 4) creates a logical linkage between Topics 3 and 5, and new references are needed to tell your story in relationship to the larger context of health systems research.

Topic 1: Define health disparities and the impact of health disparities on maternal-child health.
Topic 2: Describe public health nurse family home visiting.
Topic 3: Describe measurement of health disparities using EHR data.
Topic 4. Describe reuse of EHR data in research.
Topic 5. Describe potential to understand public health nurse home visiting intervention effectiveness relative to health disparities.

For the quality improvement activity: is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle? As with the previous example, the use of EHR data (Topic 4) in quality improvement data may be novel in the smoking cessation intervention literature and may not have been obvious in the initial literature search.

Topic 1: Describe the importance of surveillance during health care encounters to impact smoking behavior.
Topic 2: Describe change in health system policy based on evidence regarding Topic 1.
Topic 3: Describe the Quality Improvement PDSA cycle.
Topic 4. Describe the data available in the EHR to evaluate the PDSA.
Topic 5. Describe potential to understand improvement relative to the QI process using EHR data.

For the program evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program? In this example, the process of community assessment that established the need for the BMI reduction program (Topic 3) may be consistent with other community assessments and should be referenced accordingly.

Topic 1: Describe prevalence of obesity among people with diabetes.
Topic 2: Describe the BMI reduction program and proven effectiveness in reducing BMI among people with diabetes.
Topic 3: Describe the community assessment that established the need for the BMI reduction program in the community.
Topic 4. Describe the BMI reduction program evaluation as it is conducted generally.

Topic 5. Describe potential to demonstrate outcomes of the BMI reduction program for the local participants.

For the evaluation of intervention patterns associated with outcomes of adult and adolescent mothers with and without the mental health problem, he use of large datasets to examine intervention effectiveness (Topic 3) may need to be supported by new references that were not found in the original literature search.

Topic 1: Needs of high risk mothers (particularly adolescents and mothers with mental health problems).
Topic 2: Effectiveness of PHN home visiting.
Topic 3: Use of large datasets to examine intervention effectiveness.
Topic 4: Availability of PHN datasets.
Topic 5: The Omaha System.

Complete the background section using Worksheets A, B, and C, and ending with a statement of the gap as in Worksheet B. It may be necessary to update the literature matrix with additional references that support the topics that establish the foundation of the manuscript.

Read the background section to make sure the topics are in the right order and flow from general big picture of the importance of the problem to the specific gap that the project addressed. Check to make sure all topics align throughout the manuscript, from what is known and not known in the Background through alignment of the findings to the literature in the Discussion.

16.11 The Gap in Knowledge

The Background section ends with a statement of the gap in knowledge or what is not known as described in Worksheet B, followed by the purpose statement. Check the background section and purpose statement for fit with the gap in knowledge statement and adjust as needed.

For example, for the intervention effectiveness research project: are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?*

- Little is known about public health nurse home visiting interventions as related to health disparities in the outcomes of high risk mothers.

For the quality improvement activity: is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?

- Little is known about QI interventions for smoking cessation and change in smoking behavior of people served in this health system.

For the program evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?

- It is not known how the participants with diabetes and obesity in this BMI reduction program responded to the intervention.

For the evaluation of intervention patterns associated with outcomes of adult and adolescent mothers with and without the mental health problem:

- Little is known about public health nurse home visiting interventions as related to Knowledge, Behavior, and Status outcomes of adult and adolescent mothers with and without the mental health problem.

16.12 Title, Abstract, and Conclusion

Review the manuscript title, checking to make sure it describes the story. For retrospective, observational studies or any single-group intervention studies, avoid using the words "impact" or "effect" which may imply causation. Revise the title as needed to reflect the journal interests.

For example, for the intervention effectiveness research project: are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?*

- It may be most critical to announce a finding (e.g. reduction in health disparities) such as: *Health disparities of high risk mothers decreased after public health nurse home visits.*
- Alternatively, for a journal that focuses on large dataset research it may be of interest to emphasize the novel process that produced the interesting findings: *Use of large datasets to examine health disparities among high risk mothers receiving public health nurse home visiting interventions.*

For the quality improvement activity: is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?

- It may be most critical to announce a finding (e.g. reduction in smoking) such as: *Smoking decreased after implementation of a quality improvement intervention.*

- Alternatively, for a journal that focuses on PDSA cycle activities it may be of interest to emphasize the use of the PDSA cycle: *Use of the Plan-Do-Study-Act process to improve smoking cessation interventions in a large health system*.

For the program evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?

- It may be important to announce a finding (e.g. males achieve BMI reduction program goals) such as: *Male participants successful in achieving personal goals in Body Mass Index reduction program*
- Alternatively, for a journal that focuses on program evaluations it may be of interest to emphasize the program: *Body Mass Index reduction program: Male participants successful in achieving personal goals*

For the evaluation of intervention patterns associated with outcomes of adult and adolescent mothers with and without the mental health problem:

- For a journal focusing on maternal mental health it may be advantageous to describe the clinical importance of the study: *Public health nursing interventions associated with best outcomes for mothers with and without mental health problems.*
- Alternatively for a journal focusing on large dataset methods: Use of novel intervention management methods with large nursing datasets to model public health nurse home visiting interventions and outcomes.

Journals may have style requirements for title length and word usage. Choose a suitable title that describes the story in the context of the journal requirements.

16.13 Rewrite the Abstract

After telling and re-telling your story, you may find that you need to re-write the abstract to ensure alignment with title, purpose, methods, results, discussion, and conclusions. Format the abstract according to author guidelines for the selected journal, giving attention to use of a structured abstract with prescribed headings (or no headings) and number of words. Review to ensure that the abstract tells the story according to the title and all completed sections.

16.14 Write the Conclusions Section

Often, but not always, manuscripts end with a conclusion section. The conclusion is usually one paragraph that summarizes the need for the project, provides a high level overview of the findings, and ends with next steps for the project. If there is no

conclusion section within the author guidelines, this should be the last paragraph of the Discussion section. The first paragraph of the discussion and the last sentence(s) of the abstract may be used to draft the Conclusion section. Then you may revise to ensure that you are capturing the messages of your story succinctly, clearly, and in a compelling way to inspire and generate enthusiasm for your manuscript.

For example, for the intervention effectiveness research project: are public health nursing interventions associated with changes in health equity (reduced disparities in health outcomes) *among high risk mothers after public health nursing interventions?*

- Addressing health disparities is a high priority in maternal-child health nationally and globally [23–25]. In this project we examined a large dataset of public health nurse home visiting interventions to understand the associations between interventions and health equity outcomes for high risk mothers. Improved health equity was seen as health disparities decreased. More interventions were associated with larger decreases in disparities. However, health disparities between low and high risk mothers persisted, with race/ethnicity contributing to risk. These issues must be addressed to improve practice and policy in order to improve population health. Further research is needed to validate these findings in other datasets.

For the quality improvement activity: is the QI intervention associated with change in behavior of people who smoked cigarettes on admission to care after our QI PDSA cycle?

- Health care providers are known to influence smoking behavior when implementing evidence-based surveillance interventions [9]. In this PDSA QI intervention cycle smoking decreased after implementation of a quality improvement intervention. However, few people quit or cut down after only one or two encounters. During this cycle, males quit or cut down more than females, and cutting down or quitting was associated with the diagnosis of a respiratory condition. The health system is expanding use of the intervention across additional clinic settings. Additional PDSA cycles are needed to incorporate gender- and condition-specific evidence-based interventions.

For the program evaluation: Is the BMI reduction program associated with change in obesity among people with diabetes who participated in the BMI reduction program?

- The Body Mass Index (BMI) is known to predict poor health outcomes for individuals with diabetes and obesity [10]. In this evaluation of a local implementation of the BMI reduction program, people with diabetes with elevated BMI who participated in at least 80% of the program were successful in reducing BMI. In particular, male participants were most successful in achieving personal BMI reduction goals during the program. Program outcomes aligned with previous

program effectiveness reports and participants were highly satisfied with the program [10]. This program is being replicated at several sites. Further research is indicated to understand gender-related differences in the program outcomes.

For the evaluation of intervention patterns associated with outcomes of adult and adolescent mothers with and without the mental health problem:

- Mental health problems are serious and pervasive among high risk mothers, contributing to lifelong disparities in health and social problems [12–14]. In this project we examined a large dataset of PHN home visiting interventions to understand differential relationships among intervention groups for adult and adolescent mothers with and without the mental health problem. We discovered four data-driven intervention groups, showed that PHNs tailored interventions, and demonstrated differential improvement following PHN intervention with variability between groups. This research supports the importance of PHN home visiting for mothers with and without mental health problems, and the need for further research to develop new intervention approaches in order to optimize outcomes for adults with mental health problems.

16.15 Polishing Tips

Check to see that word usage is consistent throughout. Minimize use of acronyms to the minimum that are widely recognized in the field, abbreviations (e.g. vs., etc.), contractions (e.g. don't, they'd, weren't), and possessives (e.g. nurse's, intervention's). If acronyms are necessary, ensure that acronym use is defined at the beginning of the manuscript, and only the acronym is used after the first use except at the beginning of a sentence. Check to make sure that the purpose statement is the same in the abstract, the background (just after the gap statement and before the methods section), and the discussion section first paragraph. Remove adjectives whenever possible. Use a "just the facts" voice to describe the importance of the findings and throughout. This may seem to diminish the impact; however, when the facts speak for themselves, the message will be clear, straightforward, unbiased, and extremely powerful.

16.16 Styles and Author Guidelines

Manuscripts should be prepared according to the author guidelines for a particular manuscript. Other guidelines such as SQUIRE should be used as applicable [26]. Such guidelines are very helpful for defining details of the report and can be applied to the manuscript as a checklist. However, they do not take the place of careful thought and planning regarding the story line as described above.

SQUIRE is an acronym that means Standards for QUality Improvement Reporting Excellence. The SQUIRE guidelines provide standards for dissemination

of projects that evaluate improvement in the healthcare quality, safety, and value [26]. As we have seen in Chapters 1 and 2, quality improvement is a particular case among the overall perspectives on the evaluation of outcomes related to healthcare interventions. The PIO MM worksheets and process described in this book align with the intent and content of SQUIRE guidelines, and completing Worksheets A–F will position you to address the SQUIRE guideline components. This should be done after you have completed your draft manuscript in order to ensure that the story is well described, to ensure that the guideline helps to tell the story. The same may be true of author guidelines. Knowing and understanding word counts and sections specific to a particular journal is important, but telling the story well is of the utmost importance; and editing a well-told story after it is drafted is easier that trying to find the story in a sea of unrelated words, facts, and figures.

Reflection Questions

- How does the information in the worksheets reflect the story of the project?
- How does the iterative writing and reviewing process support overall alignment of the story?
- How does the SQUIRE guideline support the writing process?
- How do author guidelines relate to the writing process?
- After drafting the manuscript, has the story changed or evolved? If so, in what aspects?

References

1. Centers for Disease Control and Prevention [US] (1999) Framework for program evaluation in public health. MMWR 48(RR11):1–40
2. Chapter 3 Single group before-and-after design [Internet]. [cited 2017 May 12]. Available from: http://158.132.155.107/posh97/private/research/evaluation/Chp_3.pdf
3. Kernler D. Section 1.2: observational studies versus designed experiments.[Internet]. [cited 2017 May 12]. Available from: https://faculty.elgin.edu/dkernler/statistics/ch01/1-2.html
4. Nurse-family partnership. [Internet]. [cited 2017 May 12]. Available from: https://www.nursefamilypartnership.org/ (2014)
5. Monsen KA, Chatterjee SB, Timm JE, Poulsen JK, McNaughton DB (2015) Factors explaining variability in health literacy outcomes of public health nursing clients. Public Health Nurs 32(2):94–100
6. Monsen KA, Radosevich DM, Kerr MJ, Fulkerson JA (2011) Public health nurses tailor interventions for families at risk. Public Health Nurs 28(2):119–128
7. Monsen KA, Radosevich DM, Johnson SC, Farri O, Kerr MJ, Geppert JS (2012) Benchmark attainment by maternal and child health clients across public health nursing agencies. Public Health Nurs 29(1):11–18
8. Monsen KA, Brandt JK, Brueshoff B, Chi CL, Mathiason MA, Swenson SM, Thorson DR (2016) Social determinants and health disparities associated with outcomes of women of childbearing age who receive public health nurse home visiting services. J Obstetric Gynecol Neonatal Nursing 46(2):292–303

9. Institute of Medicine [US] Committee on Smoking Cessation in Military and Veteran Populations (2009) Combating tobacco use in military and veteran populations. In: Bondurant S, Wedge R (eds) Tobacco-control programs: evidence-based practices. National Academies Press [US], Washington, DC
10. Guh DP, Zhang W, Bansback N, Amarsi Z, Birmingham CL, Anis AH (2009) The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. BMC Public Health 9:88
11. Garcia C, McNaughton D, Radosevich DM, Brandt J, Monsen K (2013) Family home visiting outcomes for Latina mothers with and without mental health problems. Public Health Nurs 30(5):429–438
12. Torres-González F (2009) The gap in treatment of serious mental disorder in the community: a public health problem. Ment Health Fam Med 6(2):71
13. Patel V, MAJ M, Flisher AJ, Silva MJ, Koschorke M, Prince M, Tempier R, Riba M, Sanchez M, Campodonico FD, Risco L (2010) Reducing the treatment gap for mental disorders: a WPA survey. World Psychiatry 9(3):169–176
14. World Health Organization. WHO mental health gap action programme: maternal mental health. (WHO) [Internet]. [cited 2017 May 12]. Available from: http://www.who.int/mental_health/maternal-child/maternal_mental_health/en/
15. Chi C. Data mining for translation to practice. Presented at the second international conference on research methods for standard terminologies, West Paul, MN, 15 April, 2015
16. Institute of Medicine [US] (2014) Capturing social and behavioral domains and measures in electronic health records: phase 2. National Academies Press, Washington, DC
17. Monsen KA, Bekemeier B, P Newhouse R, Scutchfield FD (2012) Development of a public health nursing data infrastructure. Public Health Nurs 29(4):343–351
18. Minnesota Department of Health (MDH). Recommendations regarding the use of standard nursing terminology in Minnesota. [Internet]. [cited 2017 May 12]. Available from: http://www.health.state.mn.us/ehealth/standards/standardsdocs/nursingterminology082114.pdf
19. Dunn OJ, Clark VA (2009) Basic statistics: a primer for the biomedical sciences. Wiley, Hoboken, NJ
20. Streiner DL, Norman GR, Cairney J (2014) Health measurement scales: a practical guide to their development and use. Oxford University Press, Oxford
21. McDonald JH (2009) Handbook of biological statistics. Sparky House Publishing, Baltimore, MD
22. Martin KS (2005) The Omaha System: a key to practice, documentation, and information management, 2nd edn. Health Connections Press, Omaha, NE
23. Bryant AS, Worjoloh A, Caughey AB et al (2010) Racial/ethnic disparities in obstetric outcomes and care: prevalence and determinants. Am J Obstet Gynecol 202(4):335–343
24. Centers for Disease Control and Prevention [US]. Healthy people 2020. Maternal and child health objectives. [Internet]. [cited 2017 May 12]. Available from: https://www.healthypeople.gov/2020/topics-objectives/topic/maternal-infant-and-child-health
25. National Research Council and Institute of Medicine, Board on Children, Youth, and Families, Commission on Behavioral and Social Sciences and Education (2000) In: Shonkoff J, Philips D (eds) From neurons to neighborhoods: the science of early childhood development. National Academy Press, Washington, DC
26. SQUIRE. Promoting excellence in healthcare improvement reporting. [Internet]. [cited 2017 May 12]. Available from: http://www.squire-statement.org/

Synthesis, Next Steps, and Epilogue

17

17.1 Planning Next Steps

The scientific problem-solving method, the nursing process model, the quality improvement model, and many other models depict a spiral of assessment, planning, implementation, and evaluation, in which evaluation findings inform the next cycle's assessment and lead to further planning, implementation, and evaluation [1–6]. In this way scientific knowledge grows, health care quality improves, and evidence-based programs are revised as needed and may be extended or expanded.

This cycle is relevant to intervention effectiveness research, quality improvement activities, and program evaluation [2–6]. Intervention effectiveness researchers will recognize the literature review (assessment) to develop a study purpose and specific aims (planning), methods implementation, and interpretation of results (evaluation). Researchers typically describe next steps for additional research to further advance the science [3, 4]. Quality Improvement professionals will recognize the PDSA cycle [plan (planning) do (implementation) study (evaluation) act (apply findings)]. It is a hallmark of quality improvement to implement iterative hypothesis testing cycles quickly in order to learn how best to implement evidence-based changes within practice settings [5, 6]. Program planners and evaluators will recognize that the basic steps of the framework for program evaluation embedded within the usual routine of public health and service organizations. Every day in such settings, evaluators consult stakeholders and assess communities and populations (assessment); define program goals with guiding questions (planning); collect data about the program (implement); analyze, interpret data, and make judgments based on the data (evaluate); and share lessons learned with stakeholders [2]. In this chapter the project cycle ends and also begins again as you consider next steps.

Given that intervention effectiveness research, quality improvement activities, and program evaluation are based in known efficacious or effective interventions, care processes, or programs these studies have considerable likelihood of demonstrating positive outcomes in real world settings. Following from these next steps are mandates to disseminate findings (as described in Chapter 16), follow through

© Springer International Publishing AG 2018

K.A. Monsen, *Intervention Effectiveness Research: Quality Improvement and Program Evaluation*, DOI 10.1007/978-3-319-61246-1_17

with stakeholders (as described below), and begin plans for furthering the work through another cycle of intervention effectiveness research, quality improvement activities, or program evaluation.

For intervention effectiveness research, translation of findings into practice may be considered the next step in the pathway from bench science to efficacy to effectiveness to practice, policy, and public health improvements; this means widespread dissemination and adoption of effective interventions [3]. Many interventions that are designed to improve population health are directed to individuals; however, the availability of these interventions may vary across populations, and access to effective interventions depends on the political will of communities, health systems, health care professionals and leaders, and government agencies and decision makers [3]. For these reasons, translation of intervention effectiveness findings is variable and the health outcomes of particular populations may be suboptimal [3]. While implementation science refers to "the study of methods to promote the adoption and integration of evidence-based practices, interventions and policies into routine health care and public health settings" [4] it is possible to continue the implementation effectiveness research trajectory by extending the intervention evaluation to other populations and settings, while partnering with communities and stakeholders to advance adoption of the intervention. Key to this process is the ability to leverage positive relationships with stakeholders throughout the intervention effectiveness research cycle.

For quality improvement, further iterative cycles in additional units, organizations, and settings may be considered the next step in the pathway to spread a successful implementation process toward improved health care quality [5]. Spread is the process of advancing a successful change from an initial pilot unit or pilot populations and replicating that change across other settings. Key to successful spread is documentation of contextual factors such as infrastructure issues, task sequences, and responsiveness of stakeholders [5]. Knowledge of change theory and experience using change management skills can support successful spread of quality improvement pilot interventions through successive PDSA cycles [6].

For program evaluation, informing stakeholders of program outcomes in relationship to sample characteristics and other contextual factors, and providing recommendations upon which decisions about programs may be made may be considered the next steps after completing the program evaluation and report [2]. Using PIO MM as the basis for program evaluation will convey a comprehensive, holistic portrayal of the program outcomes as it operationalizes all aspects of the CDC logic model, and will increase the likelihood that stakeholders will see the findings as credible and useful. Stakeholders who participate in defining and operationalizing PIO MM elements and gathering data that they find credible may be more likely to be receptive to evaluation findings and likewise may be more willing to act on recommendations [2]. Using the PIO MM ensures that findings are reported within the necessary context for interpreting findings based on the level of the data. Follow up to solicit questions and feedback may also ensure that the evaluation findings are received, understood, and used correctly [2]. Follow-up may also prevent lessons-learned from being overlooked as important decisions are considered. The

evaluator should be available to stakeholders during the decision-making phase that follows the evaluation in order to aid in interpreting findings while keeping in mind all aspects of the study [2].

Keeping in mind these nuances, plan to meet with stakeholders of your future study/next cycle. Consider diverse scenarios that future findings may show, and rehearse eventual use of the findings. This will build capacity to translate new knowledge into practice or policy, and will help to prepare stakeholders for the next cycle [2].

Extending the research using PIO MM can lead to diverse next steps. These are limited only by your imagination. Consider the following questions as you plan the next steps in your intervention effectiveness research, quality improvement activities, or program evaluation project. Keep in mind that the single group before and after design may extend intervention effectiveness research, quality improvement activities, and program evaluation to other settings and populations, while other designs may be needed to analyze the efficacy of a new intervention, to define a novel concept, or to test an innovative quality improvement strategy.

17.2 Questions to Inspire Next Steps

Next steps should be incorporated within the manuscript or other reports of your project. The next steps you specified describe further intervention effectiveness research, quality improvement activities, and/or program evaluation projects that may extend the work further. Some next steps will incorporate other methods and designs. In addition to those next steps, consider the following questions and create a list of those projects or studies that are either most important next steps, or most achievable next steps.

Based on your findings:

- Did you learn about intervention effectiveness, quality improvement, or outcomes for a particular problem? Will/should the same or other problems be further evaluated?
- Was there a relationship between interventions and other variables in the PIO MM? Will/should these relationships with interventions be further explored?
- Did you discover differential findings of population subgroups within the sample? Will/should the same or other populations of interest be examined?
- You have established a precedent for using data to study intervention effectiveness, quality improvement, or program evaluation. Will/should regular evaluations be conducted using these methods in this setting or program? If so, why? What should be changed? What should stay the same?
- What are the skills and partnerships that are is needed to provide data, support methods, and champion findings? Consider stakeholders who were helpful in the project. Were there others you wished to involve? If so, consider adding them. Who will provide support for the next steps in terms of funding, data, and manpower?

Based on the current state of the translation, how is spread or translation of the intervention addressed in your recommendations for practice or policy? In practice:

- What recommendations have been adopted? If recommendations were not adopted, why not? Should they have been? What are the consequences of not adopting them? Should these be evaluated at the individual/system/community levels?
- How have the findings been incorporated into practice at the individual level? Should individual level interventions be evaluated for effectiveness/quality/outcomes? In what way?
- How have the findings been employed to improve health systems? How many units, hospitals, agencies, and health systems now use the intervention? In what way does this relate to your project and recommendations?
- How will evidence-based practice be furthered in the organization? What are the attitudes of clinicians and leaders related to the study findings and future evidence-based intervention effectiveness research, quality improvement activities, or program evaluation projects?

In policy:

- Will administrators and leaders be informed? What are their opinions about the project? What is the relationship between leadership opinions and policy regarding evidence-based practice translation?
- Will findings be used to justify positions or other funding? How does funding show the effectiveness of policy? Should funding be incorporated into an evaluation of program value?
- How have the findings been employed at the system level to support practice? Should the system level interventions be evaluated for effectiveness/quality/outcomes? How do the system level interventions relate to individual level outcomes?
- Will findings be used to advance policy discussions or support policy change in local, regional, or national jurisdictions? What are the attitudes of decision makers related to the study findings and future evidence-based intervention effectiveness research, quality improvement activities, or program evaluation projects?

Now review your list of high priority *and* most achievable next steps. What resources would be needed in the short term to complete those that are high priority and achievable? Those for with the resources are available may be the project that will be best for the short term, while building capacity and resources to extend the research for other important next steps.

17.3 Building Evidence on Evidence

Sharing evidence from intervention effectiveness research, quality improvement activities, and program evaluation is essential for building and extending the translational research agenda in healthcare. To that end, we are building a database of all

studies using the PIO MM. Everyone who completes this survey will receive a certificate of accomplishment and a personal message from the author. Go to https://www.surveymonkey.com/r/PIOMM to let us know about your work. In addition, answers to a survey regarding next steps will enable better understanding of how intervention effectiveness research, quality improvement activities, and program evaluation are improving health care quality and population health. Everyone who submits a survey response will receive an annual report of summary findings.

The survey also provides an opportunity to offer feedback for improving future editions of this book, as it is essential to make the process of intervention effectiveness research, quality improvement activities, and program evaluation more transparent and accessible to educators, students, administrators, and clinicians alike.

In planning next steps and participating in the survey, you have created a mental model of the importance of your project and its influence in practice and policy. Take a moment to write this as an epilogue (a comment on or a conclusion about your project) that you will use as your elevator speech when asked about your completed project, and what you will do next.

Reflection Questions

- How does the process of conducting a project relate to personal development as a doctoral student? A health system leader?
- What advice would you give to someone who was about to begin intervention effectiveness research, quality improvement activities, or program evaluation?
- What are your next steps following manuscript submission and publication?

References

1. Basavanthappa BT (2004) Fundamentals of nursing, 1st edn. Jaypee Brothers Medical Publishers, New Delhi
2. Centers for Disease Control and Prevention [US] (1999) Framework for program evaluation in public health. MMWR 48(RR11):1–40
3. Gonzales R, Handley MA, Ackerman S, O'Sullivan PS (2012) Increasing the translation of evidence into practice, policy, and public health improvements: a framework for training health professionals in implementation and dissemination science. Acad Med 87(3):271
4. NIH implementation science information and resources [Internet]. [cited 2017 May 12]. Available from: https://www.fic.nih.gov/researchtopics/pages/implementationscience.aspx
5. Stillwell SB, Fineout-Overholt E, Melnyk BM, Williamson KM (2010) Evidence-based practice, step by step: asking the clinical question: a key step in evidence-based practice. AJN Am J Nurs 110(3):58–61
6. Institute for Healthcare Improvement. The science of improvement: how to improve. [Internet]. [cited 2017 May 12]. Available from: http://www.ihi.org/resources/pages/howtoimprove/scienceofimprovementhowtoimprove.aspx