CBS **Study** with **PHYGITAL** Books

**Phy**sical Book + **Digital** App

Now make your learning More **Interactive, Integrated & Simplified** with this **PHYGITAL** Edition

Golden Points • Solved/Unsolved papers • MCQs • Important Topics from the book • Mini Tests/Semester Tests & Grand Tests • High Yield Images/Tables • Important Mastermind Videos • Recent Updates

**Nursing Research and Statistics** *PHYGITAL*

*Based on* **INC Syllabus 2021-22**

LOOKinside

Celebrating 50 Glorious Years of Publishing — Dedicated to Education

**BONUS** CONTENT
**15+** **INFOGRAPHIC VIDEOS** of Research and Statistics

# Nursing Research and Statistics

*As per the Revised INC Syllabus (2021-22)*

## Special Features

- Authored & Reviewed by Subject Experts
- **300+** Figures, Tables and Flowcharts covered
- **100+** Practical Examples covered extensively
- **12+** Important Annexures on reference tables
- **300+** Subjective & Objective Questions
- Extensive coverage of Research Designs
- Research Methodology and Biostatistics covered
- Perfect amalgamation of Theoretical and Practical Aspects

*Foreword*
***Latha Venkatesan***

CBSPD Dedicated to Education
Nursing Knowledge Tree
**CBS Publishers & Distributors Pvt. Ltd.**

**Sukhpal Kaur • Muthuvenkatachalam S**

# Nursing Research and Statistics

***As per the Revised INC Syllabus (2021-22)***

Nursing Knowledge Tree
An Initiative by CBS Nursing Division


**Sukhpal Kaur** PhD, MSc (N), FNRSI
*Principal*
National Institute of Nursing Education (NINE)
Postgraduate Institute of Medical Education and Research (PGIMER)
Chandigarh

**Muthuvenkatachalam Srinivasan**
PhD (INC Consortium), MSc Nursing (AIIMS, New Delhi),
D Pharm, RN (NMBA, Australia)
*Associate Professor*
College of Nursing
All India Institute of Medical Sciences (AIIMS)
Mangalagiri, Andhra Pradesh

*Foreword*
**Latha Venkatesan**


Nursing Knowledge Tree
An Initiative by CBS Nursing Division


CBSPD
Dedicated to Education

**CBS Publishers & Distributors Pvt Ltd**

• New Delhi • Bengaluru • Chennai • Kochi • Kolkata • Lucknow • Mumbai
• Hyderabad • Jharkhand • Nagpur • Patna • Pune • Uttarakhand

# Foreword

Nursing Research and Statistics play a crucial role in shaping the foundation of evidence-based nursing practices. As an educator and administrator in the field of nursing, I have recognized the pressing need for comprehensive textbooks that navigate students and practitioners through the intricacies of research methodologies and statistical analyses specific to the Indian context.

It is with great enthusiasm that I introduce the commendable work of Dr Sukhpal Kaur and Dr Muthuvenkatachalam Srinivasan, both esteemed faculty members at Institute of National Importance such as Postgraduate Institute of Medical Education and Research (PGIMER) and All India Institute of Medical Sciences (AIIMS). Their collective expertise brings a wealth of knowledge to the field of Nursing Research and Statistics. Their collaborative efforts have resulted in a textbook that stands as a beacon for aspiring nurses seeking a solid understanding of research principles.

The authors delve into research methods, statistical applications, and their relevance to nursing practice with clarity and precision. This textbook is rooted in evidence, explained with examples and tailored to meet the syllabus of the Indian Nursing Council. The simple language used in the textbook ensures ease of comprehension, making it a valuable asset for both educators and learners.

I extend my heartfelt congratulations to Dr Sukhpal Kaur and Dr Muthuvenkatachalam Srinivasan for their significant contributions to the field of nursing education and research. May this textbook pave the way for a generation of nurses well-versed in the principles of research and statistics. I wish the authors all the success.

**Latha Venkatesan**
Professor cum Principal
College of Nursing
All India Institute of Medical Sciences, New Delhi

# About the Authors

**Sukhpal Kaur** PhD, MSc (N), FNRSI is Principal at National Institute of Nursing Education (NINE), PGIMER, Chandigarh, India. She has a considerable teaching experience and a number of awards and honors to her credit. She is member of Editorial board of various national and international journals. She has published about two hundred articles in reputed journals and has contributed thirty-seven chapters to various other books. She has written seventeen books on patient education and edited eight textbooks. She has keen interest in Research, Geriatric care, Palliative care, Chronic illnesses and Women's health. She has completed eighteen research projects as principal investigator funded by DST New Delhi; DBT New Delhi; ICMR New Delhi; DST, UT Chandigarh; and PGIMER Chandigarh Research Scheme.

**Muthuvenkatachalam Srinivasan** PhD (INC Consortium), MSc Nursing (AIIMS, New Delhi), D Pharm, RN (NMBA, Australia) is currently serving as an Associate Professor at the College of Nursing, All India Institute of Medical Sciences (AIIMS), Mangalagiri, Andhra Pradesh. Previously, he held the position of Assistant Professor at the Institute of Human Behaviour and Allied Sciences (IHBAS), New Delhi. Earlier, he had also served as an Associate Professor at the College of Nursing, AIIMS, New Delhi. He obtained Doctorate in Psychiatric Nursing from INC Consortium (in collaboration with RGUHS, Karnataka) in 2017. He is a Postgraduate in Psychiatric Nursing from AIIMS, New Delhi and has also completed a Diploma in Pharmacy. He is also a Registered Nurse with the Nurses and Midwifery Board of Australia. He has published numerous research papers in national and international journals of repute. He is Deputy Editor of Indian Journal of Psychiatric Nursing. He has contributed chapters to textbooks and Training Manuals of Govt. of India. He is an executive body member of Indian Society of Psychiatric Nurses (ISPN). He received 'Best Researcher Award' from ISPN in the year 2016. Last but not least, the author attended IRON programme for Registered Nurses in Australia, International conferences in USA and Canada with travel grant from ICMR and conference organization.

# Reviewers

**A Maria Therese** PhD (N), MSc (N), MA (Hospital Administration), PGDG&C (MSN)
*Professor*
Mother Theresa Post Graduate and Research Institute of Health Sciences
Puducherry

**Anbumalar J** MSc (N)
*Vice Principal*
White Memorial College of Nursing
Attoor, Veeyanoor Post
Kanyakumari, Tamil Nadu

**A Tamil Selvi** MSc (N), PhD (N), RN, RM
*Professor cum Principal*
Mailam Nursing College
Tindivanam, Mailam, Tamil Nadu

**Anil Kumar Sharma** PhD, MSc (MSN)
*Principal*
Secretary, Nursing Research Society of India
Manikaka Topawala Institute of Nursing
Charusat, Anand, Gujarat

**Abhishek Sharma**
MSc (Mental Health Nursing)
*Principal*
Alwar Nursing College
Alwar, Rajasthan

**Anish Pathan** MSc (MSN)
*Principal*
MN College of Nursing
Bikaner, Rajasthan

**Aldrin Vas**
PhD (N), MSc Nursing (Psychiatric Nursing)
*Professor cum Principal*
Tejasvini Nursing Institute
Mangaluru, Karnataka

**Ashok Kumar**
PhD (N), MSc (N), PGDHM, FACEN
*Professor and Dean*
Atal Bihari Vajpayee University
Lucknow, Uttar Pradesh

**Amandeep Kaur**
MSc (MSN, Neurosciences) RN, RM
*Associate Professor*
Amity College of Nursing
Amity University, Gurugram, Haryana

**Ashok Kumar Yadav** MSc (N)
*Head Department of Nursing*
Sriganganagar College of Nursing
Rajasthan

**Ambili M Venugopal**
MSc (Pediatric Nursing)
*Faculty/Clinical Instructor*
College of Nursing
All India Institute of Medical Sciences
Mangalagiri, Andhra Pradesh

**Ataullah Khan** MSc (Psychiatric Nursing)
*Principal cum HOD (Nursing)*
Integral College of Nursing
Lucknow, Uttar Pradesh

*The names of the reviewers are arranged in an alphabetical order.*

**Athar Javeth**
PhD Scholar (INC Consortium) MSN
*Assistant Professor*
Mother College of Nursing
Thrissur, Kerala

**Deepak** MSc (MSN)
(Super Speciality-Gastroenterology)
*Lecturer*
College of Nursing IIMT University
Meerut, Uttar Pradesh

**Bandana Das**
MSc (Community Health Nursing)
*Principal*
Neotia Academy of Nursing
Kolkata, West Bengal

**G Dhanalakshmi**
PhD, MSc (N), MSc (Psychology)
*Professor cum HOD*
Department of Medical Surgical Nursing
Billroth College of Nursing
Chennai, Tamil Nadu

**Barnali Mukherjee**
MSc (Psychiatric Nursing)
*Professor*
Government College of Nursing
Burdwan Medical College and Hospital
Purba Bardhaman, West Bengal

**Girija Bhaskaran**
MSc (Medical Surgical Nursing)
*HOD (Department of MSN)*
Sree Balaji College of Nursing
Chennai, Tamil Nadu

**C Susila** PhD, MSc (Maternity Nursing)
*Principal*
Billroth College of Nursing
Chennai, Tamil Nadu

**HC Rawat** MSc (N)
*Principal*
University College of Nursing
Faridkot, Punjab

**Chandrani Chakrabarti**
MSc (Medical Surgical Nursing)
*Assistant Professor*
Peerless Institute of Nursing
Kolkata, West Bengal

**Harinderjeet Goyal**
PhD, MPhil, MSc (MSN), BSc (Hons.), RN, RM
*Former Principal*
Rajkumari Amrit Kaur College of Nursing
New Delhi

**Chitra**
MSc (Child Health Nursing)
*Content Strategist cum Quality Check*
Nursing Next Live

**Jacintha D'Souza** MPhil (N)
*Principal and Professor*
Father Muller College of Nursing
Mangalore, Karnataka

**Davinder Kaur** PhD (N), MSc (N)
*Professor cum Principal*
Gian Sagar College of Nursing
Ramnagar, Rajpura (Patiala), Punjab

**Jasbir Kaur** PhD, MSc (N)
*Principal*
DMCH College of Nursing
Ludhiana, Punjab



*The names of the reviewers are arranged in an alphabetical order.*

REVIEWERS

**Jayashree Ajith** MSc (N)
*Principal*
Naraina Nursing College
Lucknow, Uttar Pradesh

**Jesudian K Nath**
MSc (Mental Health Nursing)
*Professor*
United Institute of Nursing & Paramedical Sciences
Prayagraj, Uttar Pradesh

**Jyoti Sarin** PhD (N), MSc (N)
*Principal*
MM College of Nursing
Ambala, Haryana

**K Kiranmayi** MSc (N)
*Assistant Professor*
Department of Psychiatric Nursing
All India Institute of Medical Sciences
Raebareli, Uttar Pradesh

**Kirandeep Dhaliwal**
PhD (N), MSc (Mental Health Nursing)
*Principal*
MMIN, MM (DU)
Mullana, Ambala

**Lida Antony** PhD (N), MSc (N)
*Vice Principal*
Dr Moopen's Nursing College
Meppadi, Wayanad, Kerala

**Livin Varghese** MSc (Child Health Nursing)
*Principal*
Pink City College of Nursing
Jaipur, Rajasthan

**M Mumtaz** MSc (N), MPhil, MBA
*Principal*
Annai Meenakshi College of Nursing
Coimbatore, Tamil Nadu

**Mallika Vhora** PhD, MSc (MSN)
*Principal*
Muni Seva Ashram College of Nursing
Vadodara, Gujarat

**Manish Kumar Tripathi**
Pursuing PhD, MSc (Medical Surgical Nursing)
*Associate Professor*
Ramsnehi College of Nursing
Bhilwara, Rajasthan

**Manisha Gupta**
PhD (N), MSc (N), MA (Sociology), RN, RM
*Dean and Principal*
Rabindranath Tagore University Institute of Nursing
Bhopal, Madhya Pradesh

**Manju Dhandapani** PhD (INC), MSc (N)
*Associate Professor*
NINE, PGIMER, Chandigarh

**Manoj Soni** MSc (Child Health Nursing)
*Assistant Professor*
Sharda College of Nursing
Alwar, Rajasthan

**Mayengbam Benita Devi**
MSc (N), Mphil (N), (MSN)
*Principal*
Rahman Hospital College of Nursing
(Under Rahman Institute of Nursing and Paramedical Sciences)
Guwahati, Assam



*The names of the reviewers are arranged in an alphabetical order.*

## Mukesh Chandra Sharma
PhD (N), MSc (N)

*Associate Professor*
College of Nursing
AIIMS, Jodhpur, Rajasthan

## Mukesh Kumar Saini
MSc (Pediatric Nursing)

*Associate Professor*
Swami Vivekanand Nursing College
Yamuna Nagar, Haryana

## Nanak Chand Suman
MSc (Community Health Nursing)

*Associate Professor*
Government College of Nursing
BRD Medical College Gorakhpur
Uttar Pradesh

## Narendra Kumar Sharma
MSc (Mental Health Nursing)

*Principal*
Chitrini Nursing College
Sabarkantha, Gujarat

## Navneeta R Kumar
MSc (Obstetrics and Gynecology Nursing)

*Vice Principal*
Aarogyam Nursing College
Roorkee, Uttrakhand

## Neelam MSc (OBG Nursing)

*Associate Professor*
Prakash Institute of Nursing
Greater Noida, Uttar Pradesh

## Neena Veer Singh PhD (N), MSc (N)

*Lecturer*
PGI NINE Chandigarh

## Neeraj Bansal
PhD, MSc (MSN)

*Associate Professor*
Jai Institute of Nursing and Research
Gwalior, Madhya Pradesh

## Neetu Bhatnagar
MSc (Applied Psychology), (Psychiatric Nursing),
Diploma in Nursing Administration

*Principal*
Bee Enn Nursing Institute
Chak Bhalwal, Jammu

## P K Punithavathi MBBS, DGO, DMRD

*Nodal Officer*
Sri Shellapha Hospital Blood Bank

## Priya M MSc (N)

*Faculty*
College of Nursing
AIIMS, Mangalagiri, Andhra Pradesh

## R Shankar Shanmugam
MSc (N), PhD, MBA, CCHQ

*Associate Professor/Reader*
*(Nursing Research)*
College of Nursing, Madras Medical
College, Chennai, Tamil Nadu

## Rajesh Kumar Sharma
PhD (N), MSc (MSN)

*Professor and Head*
Department of Critical Care Nursing
Himalayan College of Nursing, Swami
Rama Himalayan University, Jolly Grant
Dehradun, Uttarakhand

## Rakesh Sharma
PhD (N), FNRS, MSc (N), BSc (N)

*Assistant Professor*
College of Nursing
All India Institute of Medical Sciences
Rishikesh, Uttarakhand

*The names of the reviewers are arranged in an alphabetical order.*

**Raman Kalia** MSc (N)

*Principal*
Saraswati College of Nursing
Kurali, Punjab

**Rashmi P John** PhD (Pursuing), MSc (MSN)

*Principal*
Institute of Nursing
King George's Medical University
Lucknow, Uttar Pradesh

**Rupashree** PhD (N), DGNM, PBBSc (N), MSc (N), PDF (N), RN, RM

*Associate Professor cum Vice Principal*
College of Nursing, IGIMS
Patna, Bihar

**S Kanagavalli** MSc (Psychiatric Nursing)

*Principal*
Anbarasu School of Nursing
Chennai, Tamil Nadu

**S Kanchana** PhD (N), MSc (N)

*Principal*
Omayal Achi College of Nursing
Chennai, Tamil Nadu

**Sandhya C** MSc (N)

*Faculty*
College of Nursing
AIIMS, Mangalagiri
Andhra Pradesh

**Sandhya Ghai**
PhD (N), MSc (N), MA (Pub. Admn)

*Former Principal*
National Institute of Nursing Education (NINE)
PGIMER, Chandigarh

**Sathish Rajamani**
MSc (Psychiatric Nursing)

*Professor*
DRIEMS School & College of Nursing
Cuttack, Odisha

**Shama Lohumi** PhD (Psychology), MSc (N)

*Principal*
Shivalik Institute of Nursing
Shimla, Himachal Pradesh

**Saumya Srivastava**
MSc (Oncology Nursing), RN, RM

*Nursing Tutor*
Vivekananda College of Nursing
Lucknow, Uttar Pradesh

**Sheuli Sen**
BSc, MSc (Pediatric Nursing), MBA

*Professor*
Amity College of Nursing
Amity University, Gurugram, Haryana

**Shweta Pattnaik** PhD (N), MSc (N)

*Professor and HOD*
Choithram College of Nursing
Manik Bagh, Indore, Madhya Pradesh

**Shyamala Manivannan**
PhD (N), MSc (N), RM, RN

*Former Professor and Director (Faculty of Nursing)*
Dr M G R Educational and Research Institute Deemed University
Chennai, Tamil Nadu

**Sibin K Dummini** BSc (N)

*Principal*
Jeevandeep School of Nursing
UP State Medical Faculty
Lucknow, Uttar Pradesh

*The names of the reviewers are arranged in an alphabetical order.*

**Smitha P M** MSc (N)

*Professor/HOD*
Department of Mental Health Nursing
Narayana College of Nursing,
Nellore, Andhra Pradesh

**Sonali Banerjee** PhD (N)

*Fertility Counsellor, Holistic Wellness and Yoga Mentor*
Mumbai, Maharashtra, India
Alumni: AIIMS New Delhi
NINE PGIMER, Chandigarh
MGMIHS, Navi Mumbai

**Sunil Joshi**
MSc (Community Health Nursing)

*Principal*
Tirupati College of Nursing
Pacific Medical University
Udaipur, Rajasthan

**Sushma Saini** PhD (N), MSc (N)

*Associate Professor*
National Institute of Nursing Education (NINE)
PGIMER, Chandigarh

**Thenaruvi R** MSc (MSN)

*Associate Professor*
ARCOT Srimahalakshmi
Women's College of Nursing
Vellore, Tamil Nadu

**Usha Ukande** PhD, MSc (OBG)

*Former Principal*
Choithram College of Nursing
Manik Bagh, Indore, Madhya Pradesh

**V Nirosha** MSc (N)

*Professor*
Vivekanandha College of Nursing
Elayampalayam (Post), Tiruchengode (Taluk), Namakkal, Tamil Nadu

**Vaishali Jadhav** MSc (MSN)

*Principal*
Bharati Vidyapeeth (Deemed to be University) College of Nursing
Navi Mumbai, Maharastra

**Varun Toshniwal** PhD (N)

*Professor cum Head*
College of Nursing
IIMT University, Meerut, Uttar Pradesh

**Veena V G**

*Faculty*
College of Nursing
AIIMS, Mangalagiri, Andhra Pradesh

**Vijaya Laxmi** MSc (OBG)

*Currently Director/Principal*
Sarvodaya Nursing Institute
Faridabad, Haryana

**Vipin Kumar Pillai**
MSc (Medical Surgical Nursing)

*Principal*
Udaipur College of Nursing
Udaipur, Rajasthan

*The names of the reviewers are arranged in an alphabetical order.*



REVIEWERS

# Preface

Welcome to a transformative learning journey designed exclusively for nursing students, a journey that combines the rigors of academic research with the practicality of your daily clinical life. This book is a result of our unwavering commitment to enhance your educational experience and to ensure that you are well-prepared to excel in your nursing studies.

While crafting this book, our primary goal has been to align it with the updated syllabus of the Indian Nursing Council (INC), reflecting the evolving needs and expectations of the nursing profession. We understand that research can sometimes seem daunting, filled with complex jargons and intimidating concepts. That's why we've not just created a book, but a comprehensive guide that simplifies research, offers practical approaches, and is replete with real-world examples. We have designed this book to bridge the gap between theory and practice.

One of the standout features of this book is its vibrant and colorful presentation. Real-time images, illuminating illustrations, and a student-friendly format have been carefully curated to make the learning experience engaging and visually stimulating. We have chosen to employ simple, easy-to-understand language, and to reinforce understanding, we've included flowcharts and tables, simplifying complex concepts into digestible, manageable portions.

Additionally, we have included reference tables for inferential statistics and randomization tables as a valuable bonus. These resources can be indispensable for students and educators alike, aiding in the interpretation of research data and facilitating the critical process of randomization in research studies.

As you embark on your nursing journey, this book will be your constant companion, offering insights, guidance, and support. We encourage you to explore it thoroughly, engage with the examples we have provided, and apply the knowledge you gain in your academic pursuits and clinical practice.

We extend our gratitude to the nursing community for its continuous dedication to the well-being of others. We hope that this book contributes to your professional growth and empowers you to provide best possible care to those who are in need of it.



**Sukhpal Kaur**
**Muthuvenkatachalam Srinivasan**

# CBS Nursing Knowledge Tree

## Extends its Tribute to

# Florence Nightingale

*For glorifying the role of women as nurses,*
*For holding the title of " The Lady with the Lamp,"*
*For working tirelessly for humanity–*
*Florence Nightingale will always be*
*remembered for her*
*selfless and memorable services to the*
*human race.*



**Florence Nightingale**
**(May 1820 – August 1910)**

# From Publisher's Desk

*Dear Reader,*

Nursing Education has a rich history, often characterized by traditional teaching techniques that have evolved over time. Primarily, teaching took place within classroom settings. Lectures, textbooks, and clinical rotations were the core teaching tools; and students majorly relied on textbooks by local or foreign publishers for quality education. However, today, technology has completely transformed the field of nursing education, making it an integral part of the curriculum. It has evolved to include a range of technological tools that enhance the learning experience and better prepare students for clinical practice.

As publishers, we've been contributing to the field of Medical Science, Nursing and Allied Sciences and earned the trust of many. By supporting **Indian authors**, coupled with **nursing webinars and conferences**, we have paved an easier path for aspiring nurses, empowering them to excel in national and state level exams. With this, we're not only enhancing the quality of patient care but also enabling future nurses to adapt to new challenges and innovations in the rapidly evolving world of healthcare. Following the ideology of **Bringing learning to people instead of people going for learning**, so far, we've been doing our part by:

- Developing quality content by qualified and well-versed authors
- Building a strong community of faculty and students
- Introducing a smart approach with Digital/Hybrid Books, and
- Offering simulation Nursing Procedures, etc.

Innovative teaching methodologies, such as modern-age Phygital Books, have sparked the interest of the Next-Gen students in pursuing advanced education. The enhancement of educational standards through **Omnipresent Knowledge Sharing Platforms** has further facilitated learning, bridging the gap between doctors and nurses.

At Nursing Next Live, a sister concern of CBS Publishers & Distributors, we have long recognized the immense potential within the nursing field. Our journey in innovating nursing education has allowed us to make substantial and meaningful contributions. With the vision of strengthening learning at every stage, we have introduced several plans that cater to the specific needs of the students, including but not limited to **Plan UG** for undergraduates, **Plan MSc** for postgraduate aspirants, **Plan FDP** for upskilling faculties, **SDL** for integrated learning and **Plan NP** for bridging the gap between theoretical & practical learning. Additionally, we have successfully completed seven series of our **Target High** Book in a very short period, setting a milestone in the education industry. We have been able to achieve all this just with the sole vision of laying the foundation of diversified knowledge for all. With the rise of a new generation of educated, tech-savvy individuals, we anticipate even more remarkable advancements in the coming years.

We take immense pride in our achievements and eagerly look forward to the future, brimming with new opportunities for innovation, growth and collaborations with experienced minds such as yourself who can contribute to our mission as Authors, Reviewers and/or Faculties. Together, let's foster a generation of nurses who are confident, competent, and prepared to succeed in a technology-driven healthcare system.

**Mr Bhupesh Aarora**
(Sr Vice President – Publishing & Marketing)
bhupeshaarora@cbspd.com| +91 95553 53330

# Special Features of the Book

## LEARNING OBJECTIVES

After the completion of the chapter, the readers will be able to:

- Define research.
- Describe scientific methods of acquiring knowledge.
- Describe the need and purposes of nursing research.
- Enlist the importance of nursing research.
- Describe the historical development of nursing research.
- Enlist the characteristics of a good research.

**Learning Objectives** enlist what the students will learn after studying the entire chapter.

## CHAPTER OUTLINE

- Introduction
- Definitions
- Identification of Research Problem
- Sources of Research Problem
- Criteria to Select Good Research Problem
- Formulating the Research Problem
- Elements of Problem Statement
- Identification of Variables
- Operational Definitions
- Research Objectives

**Chapter Outline** provides a quick glance of the entire chapter in one go.

## KEY TERMS

**Accidental sampling:** Selecting individuals who are readily available or easy to reach for the study.
**Accessible population:** The portion of the population that the researcher can reach out and study.
**Cluster sampling:** Dividing the population into clusters and randomly selecting entire clusters for inclusion in the sample.

**Important Key Terms** have been added in the beginning of every chapter to get a quick and easy understanding of important terms in one go.

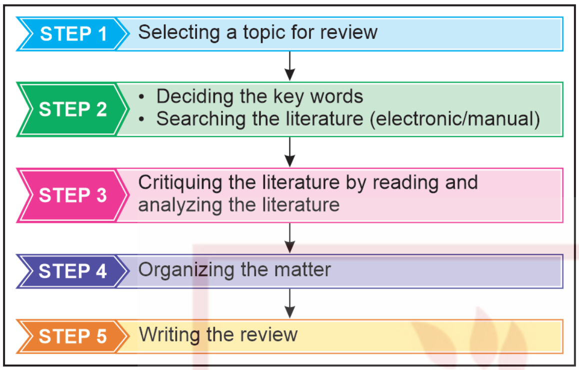


Figure 6.2: **Steps of literature review**

**Illustrations** and real-time images are used to make learning easy for students.

**Tables** have been used in Part A and B to facilitate learning in a quick way.

**Table 13.2** **References versus bibliography**

| Reference | Bibliography |
|---|---|
| References are given to the content which is quoted in the article/text or assignment | The bibliography includes list of all the sources which a researcher has gone through to conceive the idea |
| Mostly from the primary data sources | From both primary and secondary data sources |
| The list is arranged either alphabetically or numerically | The list is arranged numerically |
| Considers only in-text citations | Considers both in-text citations and other sources referred to develop an idea |
| Supports arguments by strengthening the evidence | It cannot be used in arguments |



**BOX 7.1** **Theory versus model**

Theories are general and abstract ideas that explain how things work. Models are specific and concrete ways of showing how things are or should be. Models can come from theories, but they can also come from observations or experiences. For example, one theory is that the environment affects the health of patients. One model is that nurses should consider five environmental factors (For example—fresh air, pure water, efficient drainage, cleanliness/sanitation and light/direct sunlight) in their practice

**Boxes** included in the chapters enrich students with extra information and provide them an edge over others.

**Illustration 3:**

Calculate the crude and age specific death rates of two populations A and B from the following data.

| Age group (in years) | Area A population | Deaths | Area B population | Deaths |
|---|---|---|---|---|
| Below 5 | 18,000 | 370 | 38,000 | 500 |
| 5–30 | 15,000 | 510 | 48,000 | 900 |
| Above 30 | 17,000 | 120 | 14,000 | 600 |
| **Total** | **50,000** | **1,000** | **100,000** | **2,000** |

**Solution: Crude death rate (CDR)** = Total number of deaths/Total population × 1000

**CDR (Area A)** = (1000/50000) × 1000 = **20**

**CDR (Area B)** = (2000/10000) × 1000 = **20**

Text enriched with variety of practical examples for better and quick understanding of the concepts.

Throughout the book **numerous examples** have been added from applied point of view.

**Examples of conceptual and operational definitions of 'nursing competence'**

**Conceptual definition:** Nursing competence represents the comprehensive knowledge, skills, judgment, and ethical principles that underpin a nurse's ability to provide safe, effective and compassionate patient care.

**Operational definition:** In this research study, nursing competence refers to the score obtained by registered nurses on Holistic Nurse Competence Scale (HNCS) which includes five domains such as general aptitude, staff education and management, Ethical practice, provision of nursing care and professional development as components. A higher score indicates higher levels of nursing competence.

## SUMMARY

- **Meaning and definition of research:** Research is defined as a systematic and systematic process of collecting and analyzing information to expand knowledge and solve problems.
- **Methods of acquiring knowledge:** The chapter discussed structured and unstructured methods of acquiring knowledge.
- **Problem-solving:** Research is a systematic approach to problem-solving that involves identifying, defining, and solving problems.

Important takeaway points of respective chapters have been highlighted under **Summary boxes**.

## ASSESS YOURSELF

### Long Answer Questions

2. Discuss five major sources of literature.
3. Discuss five major online sources of literature.

### Short Answer Questions

1. Define concepts.
2. Define framework.

### Multiple Choice Questions

1. A blueprint for the conduction of study is:
   a. Pilot study
   b. Research hypothesis
   c. Research design
   d. Data collection tool

Detailed **Assess Yourself** exercises in each and every chapter will facilitate structured learning and revision of the material provided in the respective chapters.

**Bibliography** allows students to learn more on nursing research and enhance their efficiency.

## BIBLIOGRAPHY


- *Abbott P. Why do we Need to Review Literature? Nurse Researcher 1993; 1(1): 14-22.*
- *Burns N, Grove S. The Practice of Nursing Research: Conduct, Critique, and Utilization.(4th ed.), 2001, Philadelphia: W.B. Saunders Company.*
- *Dempsey PA, Dempsey AD. Using Nursing Research: Process, Critical Evaluation, and Utilization. 5th ed., 2000, Philadelphia: Lippincott Williams and Wilkins.*
- *Downs FS. Hitching the Research Wagon to Theory (editorial). Nursing Research 1994; 45:195.*
- *Fisher CG, Wood KB. Introduction to and Techniques of Evidence based Medicine. Spine 2007;32 (Suppl):S66-72.*


## ANNEXURE 4: Spearman correlation coefficient table

| | α | | | | |
|---|---|---|---|---|---|
| n | 0.10 | 0.05 | 0.025 | 0.01 | 0.005 |
| 5 | 0.800 | 0.900 | 1.000 | 1.000 | – |
| 6 | 0.657 | 0.829 | 0.886 | 0.943 | 1.000 |
| 7 | 0.571 | 0.714 | 0.786 | 0.893 | 0.929 |
| 8 | 0.524 | 0.643 | 0.738 | 0.833 | 0.881 |
| 9 | 0.483 | 0.600 | 0.700 | 0.783 | 0.833 |
| 10 | 0.455 | 0.564 | 0.648 | 0.745 | 0.794 |

The **Annexures** included in the book provide extra information to students apart from their regular syllabus-based study.



SPECIAL FEATURES OF THE BOOK

# Contents

CONTENTS

## Chapter 11: Validity and Reliability of Research Tools 295–304



## Chapter 12: Analysis and Interpretation of Data 305–348



## Chapter 13: Communication, Critiquing and Utilization of Research Findings 349–392



# PART B STATISTICS

## Chapter 14: Introduction to Statistics and Organization of Data 395–417





## Chapter 15: Descriptive Statistics: Measures of Central Tendencies and Variability 418–444



## Chapter 16: Normal Distribution and Correlation Coefficient 445–460

## Chapter 17: Inferential Statistics 461–493



## Chapter 18: Vital and Health Statistics 494–509



## *Annexures 511–537*



## *Index 539–548*

that vary from individual to individual. Variables can be controlled or manipulated. They differ in the role they are given in research and in the type of measures that can be applied to them.

**Example**

**Study:** A researcher is investigating how two different methods of teaching (Video Vs demonstration) a procedure to second-year nursing students affect their performance in inserting a nasogastric tube.
**Variables:** In this study, the grade level (second year) remains the same for everyone - it's a constant condition. However, after teaching methods are applied, each student's performance in inserting NG tube is measured. Since all students will not perform identically, their scores will vary. These varying scores are what we call variables.

## Types of Variables (Fig. 4.2)

1. **Research variables:** Research variables are integral components of any scientific investigation, forming the building blocks upon which hypothesis is tested and conclusions are drawn. However, the term 'Research variable' is often used when the variables are observed in natural setting without any manipulation as in descriptive, exploratory and qualitative research. In these types of research, the primary objective is to observe and understand the variables.

**Example**

In a research study titled 'A study to assess the nutritional habits and dietary preferences of adolescent school students in a selected high school', 'nutritional habits' and 'dietary preferences' are research variables as they are observed in a natural setting without any manipulation.

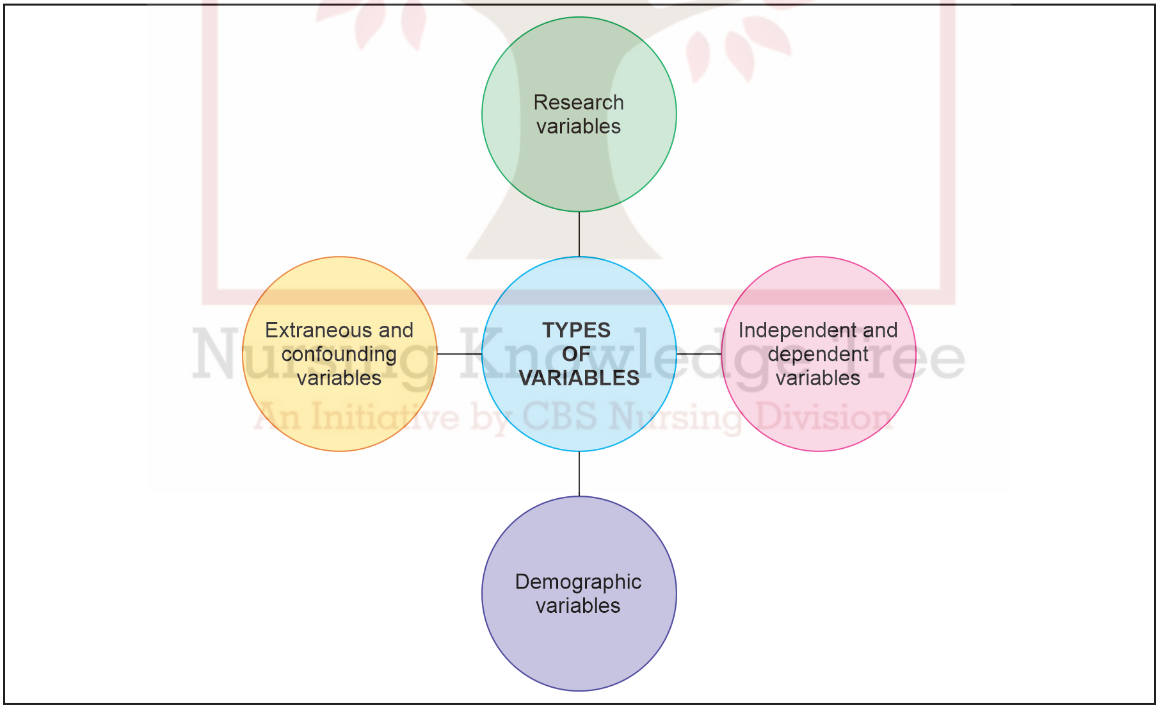


Figure 4.2: **Types of variables**

2. **Independent and dependent variables:** Researchers often examine how one variable affects another variable. In this context, the variables are classified as independent and dependent variables. These are observed in experimental and correlational research studies.
   - **Independent variable:** It is a variable that is intentionally manipulated or changed by the researcher to observe its effects on another variable. It is a variable that is believed to have an effect on the outcome of interest, which is the dependent variable. The independent variable serves as the cause or condition under investigation. Researchers change or manipulate the independent variable deliberately to observe its potential effect on the dependent variable. In controlled experimental settings, researchers assign participants to different groups, each experiencing a unique level or type of the independent variable.
   - **Dependent variable:** Dependent variable is the outcome of interest that is affected by the changes in the independent variable. It is a variable that is measured or observed by the researcher to determine if it is affected by the independent variable. This interrelationship between independent and dependent variable establishes a cause-and-effect relationship within the study.

**Example**

In a study to investigate the effect of mindful meditation on postsurgical pain among patients – who underwent abdominal surgery, the mindful meditation is the independent variable, while the postsurgical pain level reported by patients is the dependent variable.

3. **Demographic variables:** Nursing and health care research often involves a human participants. Hence, it is essential to gather comprehensive data regarding the characteristics of the study population (participants). Demographic variables are the characteristics of a population, such as age, gender, race, ethnicity, income, education, and marital status. These variables are often used in research to describe the sample population and to control confounding factors.
4. **Extraneous and confounding variables:** Extraneous and confounding variables are both factors that can affect the results of a research study (Fig. 4.3).
   - **Extraneous variables:** An extraneous variable is any variable outside the experimental design that has the potential to affect the dependent variable. Extraneous variables are not part of the study but may affect the dependent variables. They produce an association between the independent and dependent variables that is untrue, meaning the variables are not causally related.
   - **Confounding variables:** The word 'confound' means 'to confuse'. Confounding variables independently affect both the independent and dependent variables. They influence the relationship between the independent and dependent variables, making it difficult to accurately measure the true cause-and-effect relationship. Confounding variable makes it a challenge to isolate the true influence of independent variable on the dependent variable (Fig. 4.3).



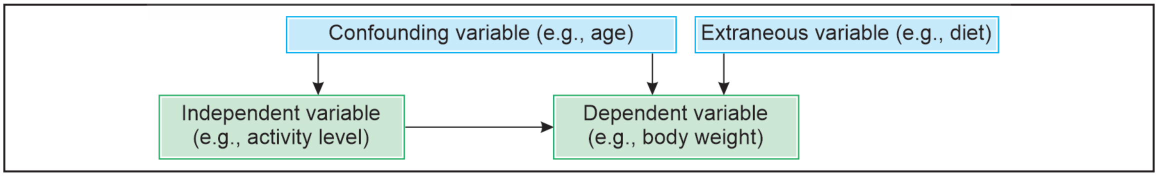


Figure 4.3: **Confounding variable versus extraneous variable**

**Example**

**Study:** 'A study to investigate the effect of activity level on body weight among adults'
**Independent variable (IV):** Activity level
**Dependent variable (DV):** body weight
**Extraneous variable:** Diet preference
**Confounding variable:** Age
**Role of extraneous variable:** Participants in the study may have varying dietary habits. Diet preference of an individual (eg. veg vs non-veg; low-calorie vs high calorie) could lead to differences in calorie intake and nutritional choices, which in turn could affect body weight (DV), irrespective of their activity levels (IV). This extraneous variable can introduce variability in body weight outcomes, complicating the interpretation of the relationship between activity level (IV) and body weight (DV).
**Role of confounding variable:** Age can be a confounding variable in this scenario as it could be related to both activity levels and body weight. Different age groups may engage in varying levels of physical activity. Younger adult participants might be more active due to higher energy levels than older adults. Age can also influence body weight due to metabolic changes with age. Due to the confounding effect of age, it becomes challenging to ascertain whether changes in body weight are due to the varying activity levels or the differences in age.

## Univariate versus Multivariate Study

- **Univariate study:** A univariate study focuses on studying a single variable. In this type of study, researchers examine the distribution, characteristics, and patterns of variation within a single variable. Univariate studies are often used to describe the characteristics of a particular phenomenon without considering relationships with other variables. For example, a study that analyzes the distribution of blood pressure readings among elderly people is a univariate study. The focus is solely on understanding the distribution of a single variable (blood pressure) without considering other factors.
- **Bivariate study:** A bivariate study involves the analysis of the relationship between two variables. Researchers explore how changes in one variable affect another variable. Bivariate studies aim to uncover associations, correlations, or cause-and-effect relationship between two variables. At least two variables are included in correlational and interventional research. For example, a correlational study to assess the relationship between physical activity and body weight is a bivariate study.
- **Multivariate study:** A multivariate study expands the analysis to include three or more variables simultaneously. Researchers investigate the complex interactions and relationships among multiple variables, accounting for potential confounding effects and interactions. For example, an exploratory study to determine the factors influencing academic performance of nursing students may include variables such as study habits, sleep patterns and socioeconomic status. This multivariate study aims to understand how these variables collectively contribute to academic outcomes.



## Quantitative versus Qualitative Variables

- **Quantitative variables:** Quantitative variables, also known as numerical variables, are those that can be measured and expressed as numerical values. These variables represent quantities with meaningful intervals between the values. They can be further categorized into two types:
  1. **Continuous variables:** Continuous variables can take any value within a specific range and can have infinite possible values. They are typically measured using instruments with fine scales. Examples include height (165.5 cm), weight (61.5 kg), temperature (98.6°F), and age.

**Table 6.2** Name and website URLs of some important nursing journals

| Name of the Journal | URL address to access |
|---|---|
| The American Journal of Nursing | https://journals.lww.com/ajnonline/pages/default.aspx |
| American Nurse | https://www.myamericannurse.com/ |
| Critical Care Nurse | https://www.wgu.edu/career-guide/health care/critical-care-nurse-career.html |
| Emergency Nurse | https://www.rasmussen.edu/degrees/nursing/blog/emergency-nursing/ |
| Geriatric Nursing | https://explorehealth careers.org/career/geriatrics/geriatric-staff-nurse/ |
| Intensive and Critical Care Nursing | https://www.sciencedirect.com/journal/intensive-and-critical-care-nursing |
| International Journal of Nursing Studies | https://www.sciencedirect.com/journal/international-journal-of-nursing-studies |
| International Nursing Review | https://onlinelibrary.wiley.com/journal/14667657 |
| Journal for Nurse Practitioners | https://www.npjournal.org/ |
| Journal of Advanced Nursing | https://onlinelibrary.wiley.com/journal/13652648 |
| Journal of Clinical Nursing | https://onlinelibrary.wiley.com/journal/13652702 |
| Journal of Nursing Education Today | https://www.sciencedirect.com/journal/nurse-education-today |
| Journal of Nursing Management | https://onlinelibrary.wiley.com/journal/13652834 |
| Nurse Education in Practice | https://www.sciencedirect.com/journal/nurse-education-in-practice |
| Nursing Outlook | https://www.nursingoutlook.org/ |
| Nursing Research | https://journals.lww.com/nursingresearchonline/pages/default.aspx |
| Oncology Nursing Forum | https://www.ons.org/onf |
| Worldviews on Evidence-Based Nursing | https://sigmapubs.onlinelibrary.wiley.com/journal/17416787 |

- **EBSCO:** EBSCO is a popular database provider that offers a wide range of academic databases and tools for researchers, librarians, and students.
- **Etoh:** Etoh, which stands for Essential Evidence Topics in Health, is a resource for evidence-based health care research.
- **Radix:** Radix is a database that offers access to nursing and allied health research materials.
- **Online journals:** The name and website URLs of some important nursing journals are provided in Table 6.2.



## STEPS OF LITERATURE REVIEW

Literature review is a systematic and organized process which involves the following steps as illustrated in Figure 6.2.

1. **Defining the research question and identifying the key terms:** The first step in conducting a literature review involves clearly defining the research question or problem statement. This question serves as the guiding principle for the review. Additionally, identifying key terms and concepts related to the research question is essential. Boolean operators such as AND, OR, and NOT can be used to combine keywords

and create more precise search queries. An example of research question and appropriate keywords are given in Table 6.3.

2. **Identifying and evaluating relevant sources:** Identifying and evaluating sources is a crucial task to ensure that the literature review includes high-quality and relevant materials. Scholarly sources, such as peer-reviewed journal articles, books, and reputable reports, should be selected. Assessing the credibility of each source involves considering author qualifications, journal status, recency of publication and methodology. Reputable journals and publishers have established quality control processes and editorial standards that enhance the credibility of the work published in their journals. It is essential to ensure that the journal has undergone peer review because peer-reviewed articles are subjected to evaluation by experts in the field, which adds a layer of quality assurance. Recently published sources (published within the last 5–10 years) should be prioritized to ensure the review incorporates the latest nursing research and practices since nursing is a fast-evolving field. The relevant literature sources related to nursing has been discussed earlier in this chapter.

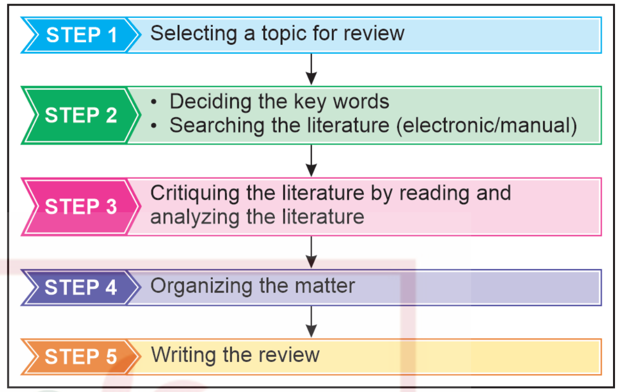


Figure 6.2: **Steps of literature review**

3. **Searching the literature:** After selecting relevant sources, the next crucial step is to identify a structured approach to searching for appropriate information. Manually searching for printed journals was the standard practice until the digital revolution and advancements transformed the literature review process, making it significantly more efficient. Here, both the manual and digital methods and some key databases used in nursing research are discussed.
   - **Manual searching of printed journals:** Manual searching of printed journals involved physically accessing and browsing through printed volumes of journals in libraries. Researchers would rely on bibliographic indices or indexes to locate specific articles. Bibliographic indices, such as the Cumulative Index to Nursing and Allied Health Literature (CINAHL), Index Medicus (Medline), and others, provided structured listings of articles, often organized by subject, author, or keyword. Researchers would consult these indices to identify relevant articles and then locate the physical journals in libraries.



**Table 6.3** **Identification of keywords from research problem statement**

| Research problem | Example of appropriate key terms for search |
|---|---|
| A study to assess the impact of telemedicine on health care access and patient outcomes in rural communities requires assessment. | • Telemedicine and Rural health care<br>• Rural health care and Health care access<br>• Telemedicine and patient outcomes<br>• Telemedicine effectiveness and rural health care |
| The effectiveness of mindfulness-based stress reduction therapy on the psychological well-being of cancer patients undergoing chemotherapy. | • Mindfulness-based stress reduction and Psychological well-being and Cancer patients<br>• Midfulness-based stress reduction and Chemotherapy |

- **Digital and online databases:** The advent of digital technology and the internet revolutionized the literature review process, making it faster and more accessible. Digital databases and search engines allowed researchers to search and retrieve a vast amount of literature from their computers. Careful and appropriate use of keywords is crucial for effective literature search using online databases.
  - **Spelling of keywords:** When searching for literature in databases, it is important to be aware of the different spellings of certain terms in British and American English. For example, the spelling of "anaemia" is "anaemia" in British, while the American spelling is "anemia". If the journal is of American origin, it may not locate the key term anaemia. So, it is recommended to use both spelling 'anemia' and 'anaemia' in the query with Boolean operator 'OR' (anaemia OR anemia). This will locate the article if it has the keyword either 'anemia' or 'anaemia'.
  - **MESH terms:** MESH (Medical Subject Headings) terms are a controlled vocabulary used to index biomedical and life sciences articles. Using MESH terms enhances the precision of the literature search. The MESH Browser is a dedicated tool provided by the National Library of Medicine (NLM) for exploring MESH terms. MESH terms in search can be incorporated while searching articles in PubMed.
  - **Boolean operators:** Boolean operators such as AND, OR, and NOT can be used to combine keywords and create more precise search queries. By strategically using Boolean operators and keywords, the researchers can efficiently find relevant literature. For example (Fig. 6.3),
    **AND:** Use AND to retrieve articles that contain both keywords. For example, "anaemia AND treatment" will retrieve article that contain both of these keywords.
    **OR:** Use OR to retrieve articles that contain either or both the keywords. For example, "anaemia OR iron deficiency" will retrieve articles that contain either 'anaemia' or 'iron deficiency' or both the terms.
    **NOT:** Use NOT to exclude keywords from the search results. For example, "anaemia NOT sickle cell" will retrieve articles that contain the keyword "anaemia" but not the keyword "sickle cell".
  - **Truncation to broaden search:** Truncation is a technique that allows the researcher to broaden the search by including variations of a keyword. To truncate a keyword, a wildcard character * is used at the end of the keyword. For example, truncating the keyword "Nurse" to "Nurs*" will return results that contain the following keywords: nurse, nursing, nurses, nursing care, etc.
  - **Field tags for narrowing search:** The strategic use of "field tags" is a valuable technique to enhance search accuracy and reduce the likelihood of irrelevant search results. Field tags are specialized operators or codes used within database search queries to designate a particular field



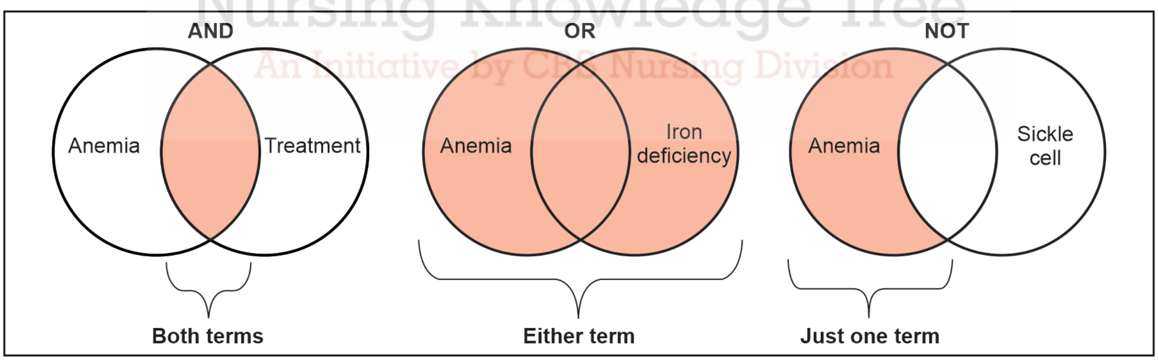


Figure 6.3: **Boolean operators and their purpose**

**Table 8.3** Quantitative research designs

| Type | | | Description |
|---|---|---|---|
| Experimental | True experimental | Pretest post-test control group design | Participants undergo a pretest, receive an experimental treatment, and then take a post-test. The control group provides a baseline for comparison. |
| | | Post-test only control group design | Similar to the pretest post-test design, but without the pretest. This design is effective in situations where a pretest may influence the study outcomes. |
| | | Solomon four, group design | Combines elements of pretest post-test design and post-test design only control group designs. Includes two experimental and two control groups to examine the effects of pretesting on study outcomes. |
| | | Factorial design | Multiple independent variables are manipulated simultaneously to understand their individual and combined effects on the dependent variable. |
| | | Randomized block design | Participants are grouped into blocks based on specific characteristics and then randomly assigned to different experimental conditions within those blocks. Helps control for confounding variables. |
| | | Cross over design | Participants experience multiple experimental conditions in a random order, with a washout period between conditions. Common in clinical trials, particularly when testing the effectiveness of different treatments. |
| | | Parallel group or Cluster-Randomized Trials (GRTs) | Group of participants or clusters are randomly assigned to different interventions. Practical when individual randomization is challenging. |
| | | Stepped Wedge Cluster-Randomized Trial | Interventions are sequentially introduced to different clusters overtime, allowing researchers to assess effects gradually. |
| | | Matching samples designs | Participants are matched based on specific characteristics before being assigned to different groups. Helps control for confounding variables. |
| | | Latin square design | Commonly used in experimental settings where researchers manipulate multiple factors, ensuring each level of each factor appears once in every row and column. |
| | Quasi-experimental Design | Nonrandomized control group | Participants are assigned to groups without randomization, useful when randomization is impractical or ethically challenging. |
| | | Time series designs | Data is collected at multiple of time before, during, and after an intervention, allowing researchers to observe trends and changes overtime. |

*Contd...*

| Type | | | Description |
|---|---|---|---|
| | Pre-experimental design | One-shot case study | A single group is observed after an intervention, providing a snapshot of outcomes without a comparison group. |
| | | One group pretest-post-test design | Participants undergo a pretest, receive an intervention, and then take a post-test. A straightforward design is lacking a control group for comparison. |
| Non-experimental design | Descriptive | Cross-sectional | Data is collected from participants at a single point in time, providing a snapshot of a population's characteristics. |
| | | Longitudinal | Data is collected from the same participants over an extended period, allowing researchers to examine changes and developments overtime. |
| | Analytical | Retrospective (Case-control) | Involves comparing individuals with a specific outcome (cases) to those without the outcome (controls) based on past events or conditions. |
| | | Prospective (cohort) | Participants are followed overtime to assess the development of a particular outcome. Useful for studying the incidence of diseases or events. |

**Table 8.4** Steps followed in conducting experimental research

| Steps | Example |
|---|---|
| STEP 1 Identify the problem | • Increased incidences of pressure ulcers are observed in neuro ICU. |
| STEP 2 Formulate research questions that need to be answered | • Whether application of preventive silicone dressings to the sacrum and heels reduces the incidences of pressure ulcers? |
| STEP 3 Formulate hypothesis | • There will be less incidence of pressure ulcer among patients who received preventive silicone dressings than that of patient who received routine care. |
| STEP 4 Select sample | • Patients admitted in neuro ICU are selected for the study using total enumeration sampling method. |
| STEP 5 Allocate sample into experimental and control group (if applicable) | • Selected sample (patients) is randomly allocated into control (routine care) and experimental (silicone dressings + routine care) group. |
| STEP 6 Gather data | • Incidence of pressure ulcer is recorded over next 6 months. |
| STEP 7 Analyze the data to test hypothesis | • The data is analyzed to test the hypothesis. |
| STEP 8 Interpretation of results | • The application of silicone dressings, in addition to routine care, in neuro intensive care unit patients is effective in preventing pressure ulcers at the heels and sacrum. |

The experimental research designs are further classified into:

- True experimental research designs
- Quasi-experimental research designs
- Pre-experimental research designs

### True Experimental Research Designs

True experimental research design is the most accurate form of experimental design. This method is often termed Randomized Controlled Trial (RCT) in biomedical research. In this method, the researchers have complete control over the extraneous variables. Hence, they can conclude that the change in dependent variable is due to the effect of independent variable.

The true experimental approach is the most powerful design for testing hypothesis of causal relationship among the variables. The primary purpose of true experimental research is to investigate the effect of an independent variable(s) on the dependent variables(s) to determine the existence of a cause-and-effect relationships between the variables. The basic structure of true experimental research design has been illustrated in Figure 8.3. This may vary from design to design as the hybrid designs use combination of basic designs to improve the control over the variables.

**Essential characteristics of true experimental design:** The **three essential characteristics** of an experimental design are manipulation, control and randomization.

1. **Manipulation:** Experimental manipulation is a process by which researchers intentionally change, modify, or influence the independent variable. Manipulation is a conscious and purposeful act by the researcher to modify the independent variable (cause) and observe the effect of manipulation on the dependent variable (effect). In other words, manipulation of a variable occurs when the researcher controls the movement of these variables.

   **Example 1:** A researcher is conducting a study to examine the effect of application of cold pack on the skin before injection of vaccine on pain level at the time of injection. Providing cold pack application to children in experimental group and withholding for children in control group are considered manipulations of independent variables. The effect of the manipulated independent variable (cold pack application) is observed on the dependent variable (pain level at the time of injection).

   *Independent variable: Application of cold pack*

   *Dependant variable: Pain level at the time of vaccination injection.*

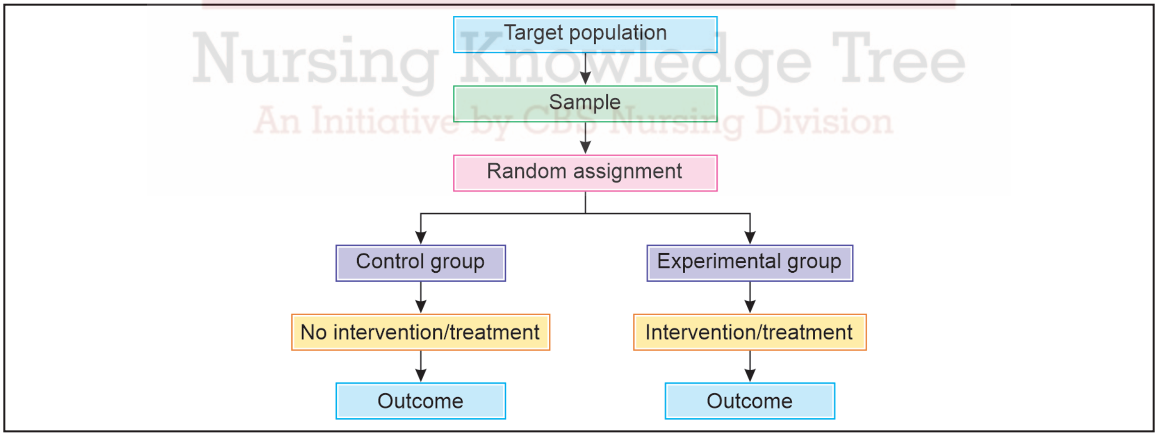




Figure 8.3: **Basic structure of true experimental research design/RCTs**

**Table 8.6** Type of blind trials

| Type of blind trial | Awareness about treatment of stakeholders involved in research study | | |
|---|---|---|---|
| | Participants | Researchers | Data analysis |
| Open trial | Aware | Aware | Aware |
| Single blind trial | Blinded | Aware | Aware |
| Double blind trial | Blinded | Blinded | Blinded |
| Triple blind trial | Blinded | Blinded | Blinded |

Table 8.7 describes different types of bias along with description and intervention.

**Table 8.7** Types of bias in experimental research along with description and intervention

| Types of bias | Description | Intervention |
|---|---|---|
| Selection bias | Systematic differences in baseline characteristics between groups (lack of homogeneity) due to allocation without randomization | • Random allocation of subjects<br>• Allocation concealment |
| Response bias | The tendency of study participants to give inaccurate, even false answers to self-reported questionnaire. | A well-structured, neutrally-worded questionnaire minimizes the response bias |
| Recall bias | Caused by inaccurate and incomplete recollection of events by study participants. | • Reduce recall period as much as possible<br>• Choosing appropriate data collection method |
| Performance bias | It happens when one group of subjects in an experiment gets more attention and care from researchers than another group. | Blinding of researchers |
| Detection bias/ Ascertainment bias | Systematic bias in the assessment of outcome measures by researchers or study participants for one group of subjects. | Blinding of outcome assessors |
| Attrition bias | Withdrawal of participants from study results in attrition bias. Less interested or sick patients are more likely to withdraw from study, thus introducing bias in outcome. | • Intensive follow-up<br>• Intention-to-treat analysis. |
| Reporting bias | Publication bias: The journals are more likely to publish positive results than negative results.<br>Outcome report bias: It occurs when researchers intentionally report only a subset of original recorded outcomes. | The funnel plot method is used to assess the potential role of publication bias. |

**Blinding:** Blinding or 'masking' is the process that prevents parties (participant or researcher or both) involved in a research study from knowing which treatment was given to which group. Blinding is an act of withholding information about the treatment assignment from people involved in the trial from the time of group assignment until the experiment is complete. Blinding is used in RCT to minimize detection and performance bias. The following are the types of blind trials.

- **Open trial:** There is no blinding process in open trial. The research participants, researchers and the data analysts are aware of the treatment allocation details; hence, there is a chance of performance and detection bias in this type of trial.
- **Single-blind trial:** The researcher knows which treatment or intervention the participant is receiving but the participants do not know whether they are receiving the intervention or placebo. This will control the placebo effect.
- **Double-blind trial:** In this method, neither the researcher nor the study participants are informed about the treatment details. This method will prevent both placebo effect and performance effect. Therefore, double-blind trials are preferred method of clinical trials.
- **Trible-blind trial:** In this method, the treatment allocation details are withheld from participants, the researcher and the individuals who measure the outcome and analyze the data. Therefore, placebo bias, measurement bias and performance bias are eliminated.

**Use of placebo in true experimental research:** Some people feel better just from the perception of receiving new treatment; this is known as the placebo effect. Due to the positive perception, the participants in experimental group may produce some placebo effect. So, researchers use placebos in control group to determine if any difference between groups are due to the new treatment or the participants' perception (placebo effect). Placebo is an inactive substance or intervention that looks similar and given the same way as the treatment/intervention being tested. Placebo is commonly used in clinical trials. However, it is challenging task to design a placebo to match with nursing interventions. So, placebo controlled RCTs in nursing research are uncommon. The new interventions are often compared with the existing routine care practices.

### Types of True Experimental Research Designs

The various true experimental research designs include:

1. Pretest post-test control group design
2. Post-test only control group design
3. Solomon four-group design
4. Factorial design
5. Randomized block design
6. Cross over design
7. Parallel Group or Cluster-Randomized Trials (CRTs)
8. Stepped wedge cluster randomized trial
9. Matching samples designs
10. Latin square design



**BOX 8.2 Symbols used in experimental designs**

- R: Random allocation to the groups
- O: Observation or measurements
- E: Experimental group
- C: Control group
- X: Exposure of the group to the experimental variable (Intervention)

These have been elaborated as follows. All the experimental designs are shown diagrammatically using certain symbols. This method was developed by Campbell and Stanley (1966). These symbols are shown in Box 8.2.

1. **Pretest post-test control group design:** This design uses a control group to determine whether the treatment/intervention makes a difference. This design has a control group, intervention, pretest and post-test. This is also known as classic experimental design. In this design, subjects are randomly assigned to either the experimental or control group. Following the pretest in both the groups, the treatment is carried out on experimental group only. After treatment, observation of dependent variable is made in both the groups to examine the effect of the manipulation of independent variable, i.e., the intervention on dependent variable. This design can be presented as shown in Figure 8.6.

   The schematic presentation of the pretest post-test control group design has been presented in Figure 8.7. The schematic representation of the design will be used henceforth to illustrate the research design.

   For example, "A study on the effect of an orientation program on the anxiety level of patients undergoing radiotherapy for first time" (Harpreet et al. 2014).

   The patients were randomly assigned into intervention and control groups by computer generated randomization. In intervention group, orientation program was used as an intervention for the patients before they underwent radiotherapy. However, in control group, routine care was provided. Anxiety was

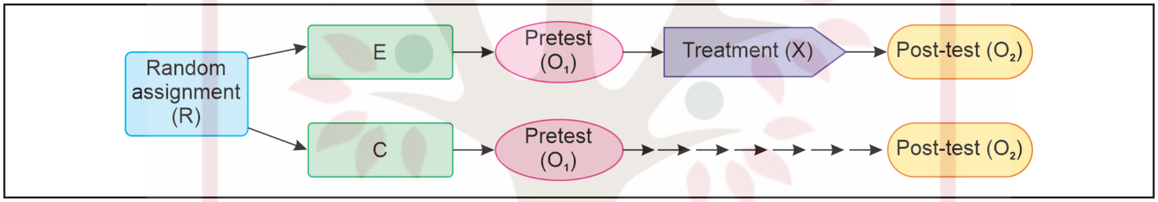


Figure 8.6: **Pretest post-test control group design**

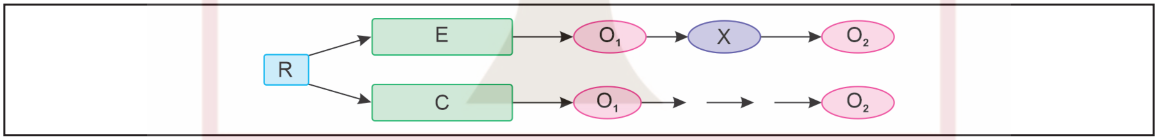


Figure 8.7: **Schematic presentation of pretest post-test control group design**



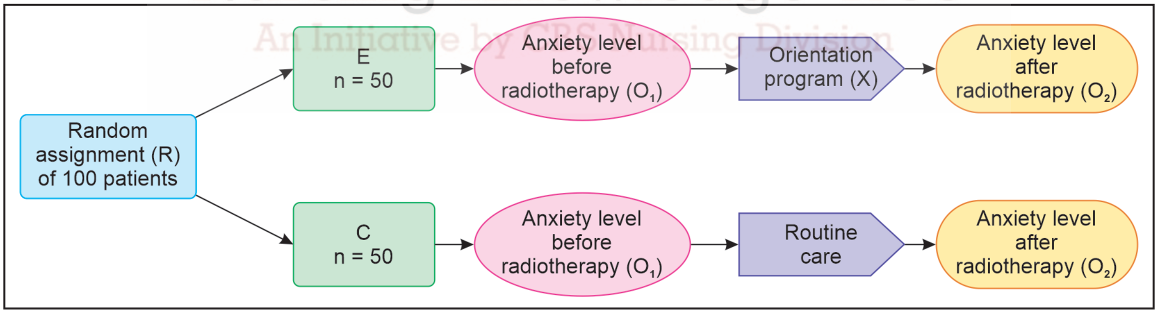


Figure 8.8: **Schematic diagram of pretest post-test control group design**

assessed in both the groups, before and immediate after radiotherapy using a 4-point, seventeen items anxiety assessment scale. Figure 8.8 illustrates the above study with a schematic diagram.

2. **Post-test only control group design (Fig. 8.9):** This design is often used in situations when a pretest of the subjects is not possible or desirable. It is composed of two randomly assigned groups - experimental and control groups. Both the groups are not tested before the introduction of an intervention. The treatment is implemented in the experimental group only. Post-test observations are made in both the groups.

**Example 1:** To study the effect of abdominal massage on bowel movements among patients diagnosed with postoperative constipation.

**Example 2:** To study the effect of negative pressure wound therapy for surgical wounds on wound healing by primary closure.

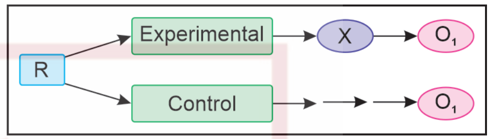


Figure 8.9: **Post-test only control design**

In both above examples, pretestsing is not possible, so post-test only design is appropriate to test the effectiveness of the intervention.

3. **Solomon four-group design (Fig. 8.10):** Solomon four-group design is a hybrid research design which is formed by combining the features of the pretest post-test design and the post-test only design. As the name suggests, the Solomon four group design has four groups: two experimental groups and two control groups. One control and experimental group undergoes pretest and post-test and another control and experimental group undergoes only post-test. Participants are randomly assigned to one of the four following groups.

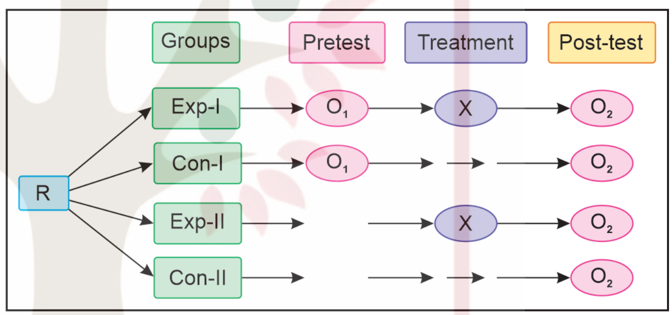


Figure 8.10: **Schematic diagram of Solomon four-group design**

1. Experimental group-I with pretest and post-test assessment
2. Experimental group-II with only post-test assessment
3. Control group-I with pretest and post-test assessment
4. Control group-II with post-test only assessment

It is illustrated as follows:

Only the participants in Experimental group-I and Experimental group-II receive the intervention/ treatment. Participants in Experimental group-I and Control group-I receive pretest and post-test, whereas participants in Experimental group-II and Control group-II receive only post-test. The outcome (dependent variable) is assessed simultaneously across all the four groups.

The Solomon four-group design has a distinct advantage of controlling threat to both the internal validity (cause-and-effect relationship) and external validity (generalizability). This design was developed to control threats to internal validity such as bias and confounders and external validity such as pretest sensitization or reactive effect which is not controlled by a standard pretest post-test control group design.

For example, in a study to assess the effectiveness of a tobacco cessation program among young adults, introduction of pretest may sensitize the participants and make them more responsive to the tobacco cessation program. However, the study results are meant to be generalized to the general population who have not received any pretest. So, the effect of intervention (tobacco cessation program) on the general population will be different than the one found in the study. This leads to poor generalizability of the

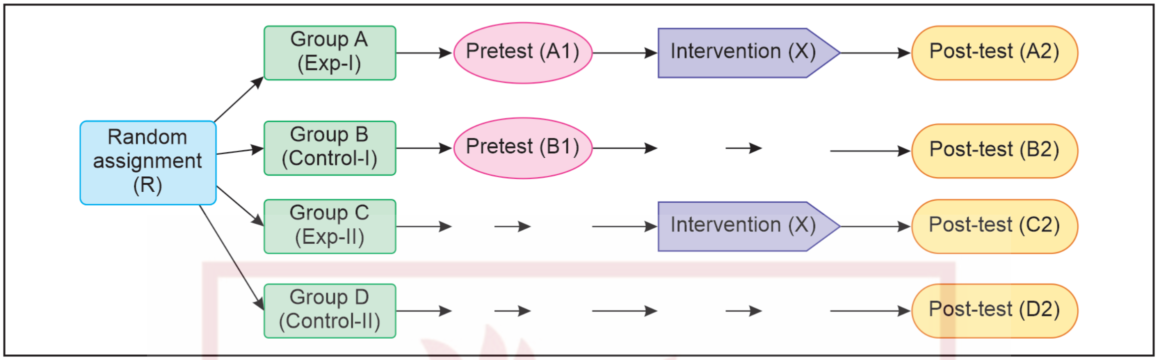


Figure 8.11: **Solomon four group design**

study findings. Solomon four-group design controls the threat to external validity by determining the pretest effect on the outcome.

The following example is used to explain how the Solomon four-group design compares the result. Groups named as A and C are experimental groups; B and D are control groups (Fig. 8.11).

The pretest sensitivity or reactive effect is assessed in three steps:

**Step 1:** The difference between A2 and C2 = the effect of pretest on the post-test in the presence of intervention. The difference between A2 and C2 determines the interaction effect of pretest on the intervention. If A2 and C2 differ significantly, it means that the pretest had an effect on the intervention and that the experiment is flawed (lacks generalizability).

**Step 2:** The difference between B2 and D2 determines the effect of the pretest on the post-test independently (without intervention). If B2 and C2 differ significantly, it means the pretesting has influenced the post-test (lacks internal validity).

**Step 3:** Compare the result of step 1 and step 2 (i.e., difference between A2-C2 and B2-D2). A difference between quantities of step 1 and step 2 means that the effect of the pretest on the post-test is different in presence of intervention, i.e. there is an interaction between pretest and intervention.

Solomon four-group design is one of the benchmarks in sociological and educational research. This design can be effectively used in nursing education research studies to control the pretest sensitivity, but unfortunately, it is less commonly used because of the need for larger sample size and complex statistical analysis methods.

Example, Chang et al. (2014) 21 conducted an RCT to assess the effect of an asynchronous e-learning curriculum on the knowledge level of medical residents. The Solomon four-group deign was used. Block randomization was done to assign each participant to any of the four groups. The group A received the pretest followed by access to the e-learning modules which was followed by the post-test. While group B was pretested at beginning of rotation and post-tested at end without receiving any access to e-learning module. The group C was given access to e-learning and post-test was done, however, pretesting was not done. Lastly, group D was post-tested only, there was neither pretest nor intervention. The results of the study revealed no pretest sensitization while use of e-learning modules improved the knowledge significantly.

4. **Factorial design:** Factorial design involves measuring the effect of two or more independent variables on a dependent variable. In other words, the researcher manipulates two or more independent variables simultaneously to observe their effects on the dependent variables. This design is particularly useful when there are more than two independent variables to be tested.

For example, a researcher wants to test the effectiveness of two psychosocial therapies and two pharmacological interventions on tobacco cessation amongst tobacco users. In this example, we have two factors: Type of psychosocial therapy and type of pharmacological therapy.

**Factor 1 (psychosocial intervention):** Two levels—Brief intervention ($\alpha 1$) and motivational counseling ($\alpha 2$)

**Factor 2 (Pharmacological intervention):** Two levels—nicotine gum ($\beta 1$) and nicotine patch ($\beta 2$).

There will be four cells to which the subjects will be randomly assigned. This is called 2 × 2 factorial design because there are two factors with two levels of treatment (Fig. 8.12). Data is collected for all possible combinations of the levels of two factors of interest. For example, participants assigned to the $\alpha 1\beta 1$ cell (Fig. 8.12) will receive brief intervention and nicotine gum therapy.

| Type of pharmacological therapy | Type of psychosocial therapy | |
|---|---|---|
| | Brief intervention ($\beta 1$) | Motivational counseling ($\beta 2$) |
| Nicotine gum ($\alpha 1$) | $\alpha 1\beta 1$ | $\alpha 1\beta 2$ |
| Nicotine patch ($\alpha 2$) | $\alpha 2\beta 1$ | $\alpha 2\beta 2$ |

Figure 8.12: **Factorial design**

Figure 8.12 depicts a 2 × 2 factorial design for the above example which can help the researcher to test multiple hypotheses, by manipulating more than one independent variable. In this study, the subjects shall be assigned to one of the four arms:

i. The first group ($\alpha 1$ $\beta 1$) will receive nicotine gum and brief intervention.
ii. The second group ($\alpha 1$ $\beta 2$) will receive nicotine gum with motivational counseling.
iii. The third group ($\alpha 2$ $\beta 1$) will receive nicotine patch with brief intervention.
iv. The fourth group ($\alpha 2$ $\beta 2$) will receive nicotine patch with motivational counseling.

The researcher shall test various research questions with this design:

- Is nicotine gum beneficial in tobacco cessation than nicotine patch?
- Is brief intervention more effective than motivational counseling in tobacco cessation?
- Is brief intervention with nicotine gum effective in tobacco cessation compared to brief intervention with nicotine patch?
- Is the motivational counseling with nicotine gum effective in tobacco cessation compared to motivational counseling with nicotine patch?

Thus, the design shall test the main effect of using a nicotine gum and nicotine patch—the main effect of the type of counseling used. The superiority of the design lies in the fact that it shall assess the interaction effects of type of pharmacological therapy and type of psychosocial therapy used.

5. **Randomized block design:** This is a simple method to reduce the variability among the treatment groups by a more homogeneous combination of the subjects through randomized block design. Randomized block design divides study participants into groups or blocks (subgroups) of equal size based on certain characteristics, and then randomly assigns the treatment to participants in each block. Randomized block design is used to block the effect of an uninterested variable (confounding variable) on the outcome variable, e.g., male and female. This is particularly useful method to control confounding variables when the sample size is small because if the subjects are allocated randomly using simple randomization without using randomized block design, then there is high probability that the confounding variable (gender) gets unevenly distributed among treatment groups.

**Example 1:** A researcher wants to study the effectiveness of a physical training program on activities of daily living among dementia patients. The physical capability varies between male and female gender

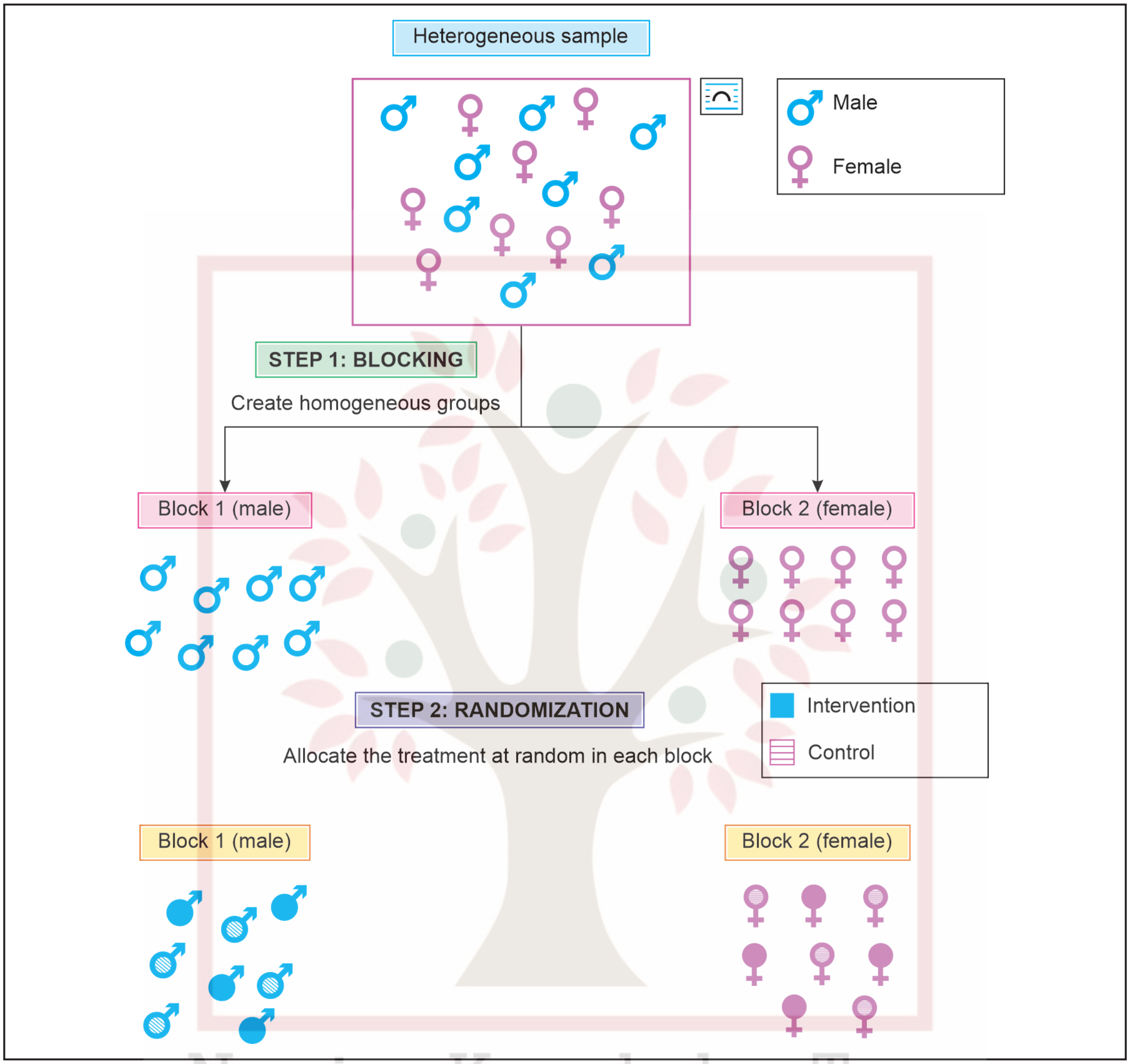


Figure 8.13: **Steps of randomized block design (Example 1)**



which may affect the outcome of the study. So, the researcher decides to create blocks based on gender to control its effect on outcome variable. The participants are divided into blocks of male and female and then the treatment (physical training program) is allocated randomly within each block (Fig. 8.13).

**Example 2:** A researcher wants to study the effectiveness of two different massage therapies (massage with kinesthetic stimulation vs massage therapy alone) on weight gain in preterm infants. In this example, the gestational age of preterm babies varies to a great extent which may influence the outcome. So, the researcher decides to divide the preterm infants to blocks (subgroups) based on gestational age: 27–29 weeks, 30–32 weeks and 33–35 weeks. Here, the gestational age is the blocking variable. Then, the intervention (massage with kinesthetic stimulation or massage therapy alone) is randomly allocated to the infants within each block (Fig. 8.14).

surveys were carried out to provide essential data on health and family welfare to the Ministry of Health and Family Welfare and other agencies for policy making and program purposes and to provide information on important emerging health issues and its sociodemographic correlates.

3. **Exploratory research design:** Exploratory research design is used to investigate a problem or phenomena that is not clearly defined. It is conducted for the purpose of better understanding of the existing problem. Exploratory studies are primarily qualitative as they are often conducted with open-ended questions to understand the phenomena. However, it is also being conducted in quantitative forms using standardized scales. For example, Patra BN et al., has conducted an exploratory study on self-stigma in patients with major depressive disorder. The stigma of the depressed patients was assessed using the stigma, the Discrimination and Stigma Scale -12.

**Types of Longitudinal or Follow-up Studies**

**Cohort Study:** Cohort study is a particular form of longitudinal study that takes cohort (which shares defining characteristics) as a sample. It is a type of epidemiological study in which, the researcher selects a group of exposed individuals and a group of unexposed individuals in a nonrandom manner, from a study population, which is initially disease free. After the selection, both groups are followed to compare the incidence of the disease. If a positive association exists between the exposure and the disease, we expect the incidence of disease in the exposed group to be greater than the incidence of disease in the unexposed group.

There are two types of cohort studies: prospective and retrospective (or historical) cohorts. Prospective cohort studies follow a cohort into the future for a health outcome (Fig. 8.31). Retrospective cohort studies track the cohort back in time for exposure information after the disease/outcome has occurred (Fig. 8.32). Both types of cohort studies are also known as longitudinal or follow-up studies (Table 8.10).

**Example:** Suppose we want to study the association between levels of physical activity and coronary artery disease (CAD) (Table 8.11). We defined a cohort, measured the levels of physical activity (exposure) in the population and divided the population into two groups based on the levels of physical activity into low and high levels. Among the cohort, some developed CAD while some did not develop the disease.

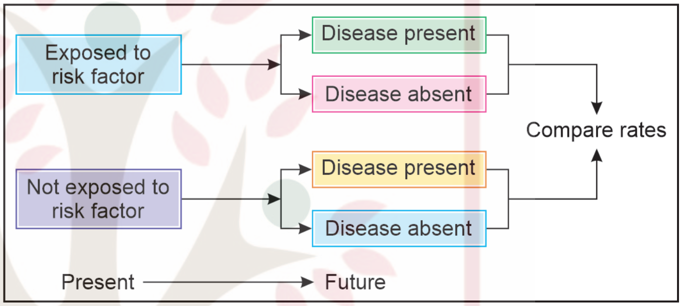


Figure 8.31: **Schematic presentation of prospective cohort study design**

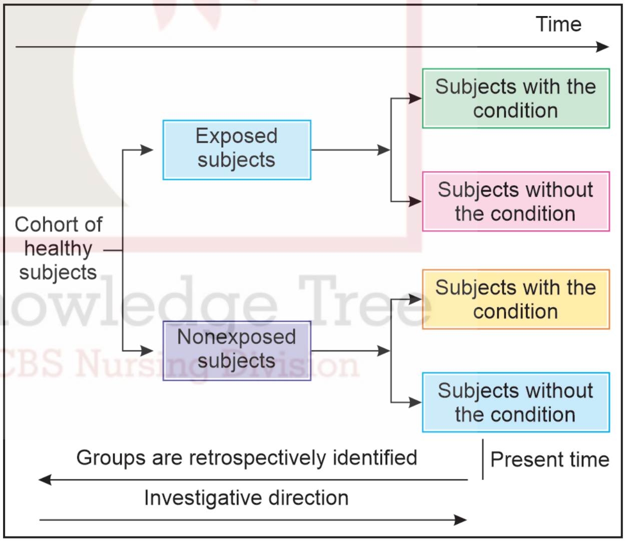


Figure 8.32: **Schematic presentation of retrospective cohort study design**

**Table 8.10** Cohort study design

The observed data will be tabulated as follows to calculate incidence rate and risk ratio.

| | Cases (with disease) | Controls (without disease) | Total |
|---|---|---|---|
| Exposed | a | b | a + b |
| Not exposed | c | d | c + d |

a – exposed and developed disease
b – exposed and not developed disease
c – not exposed and developed disease
d – not exposed and not developed disease
Incidence of disease in exposed (proportions are those at risk) = a/(a + b)
Incidence of disease in unexposed (proportions of those at risk) = c/(c + d)
Risk ratio is defined as the ratio of the risk of developing disease in the exposed group to the risk of developing disease in the unexposed group.

$$\text{Risk ratio} = \frac{a/(a+b)}{c/(c+d)}$$

**Table 8.11** Cohort design to study the association between levels of physical activity and coronary artery disease (CAD)

| Levels | Developed CAD | Did not develop CAD | Total |
|---|---|---|---|
| Low levels of physical activity | 84 | 2916 | 3000 |
| High levels of physical activity | 87 | 6913 | 7000 |

The proportion of exposed (low levels of physical activity) who developed CAD is 84/3000. The proportion of the diseased (CAD) among unexposed (high levels of physical activity) is 87/7000. Accordingly, we see that proportion of CAD in exposed was higher than the proportion of the diseased among unexposed.

Incidence of CAD in exposed (proportions are those at risk) = 84/3000 = 0.028

Incidence of disease in unexposed (proportions of those at risk) = 87/7000 = 0.012

$$\text{Risk ratio} = \frac{0.028}{0.012} = 2.33$$

**Interpretation:** A risk ratio greater than 1.0 indicates a positive association, or increased risk for developing a disease. Risk ratio 2.33 means the exposed group (low levels of physical activity) has 2.33 times the risk of having CAD compared to the unexposed (high levels of physical activity).



**Potential biases in a cohort study**

**Bias in assessment of outcome:** If the person who is assessing the outcome is aware of the exposure status of the subject and the hypothesis being tested, then he may be biased in judging the disease or outcome status. This problem can be addressed by masking the person who is carrying out the outcome assessment to the exposure status of the subjects. Method of assessing the outcome (such as investigation to diagnose a disease) can also introduce bias in the study.

Biases from nonresponse and loss to follow up: Loss to follow up can be a serious problem in a prospective study.

### Retrospective or Case-Control Study Design

Retrospective (Retro = looking at the past) or a case control study design means to start from a disease or outcome in the present and going back in time to link it with a presumed cause or exposure in the past.

Research direction is backward in time. Data is collected through record reviews or asking subjects about the information exposure, behavior or instances which occurred in the past that will explain the research question or hypothesis. To understand further, suppose we need to assess the relationship between the exposure and a disease, a group of subjects with disease or outcome (cases) and another group of subjects without the disease or outcome (controls) are selected. Then we determine the proportion of cases and controls who were exposed and what proportion were not (Fig. 8.33).

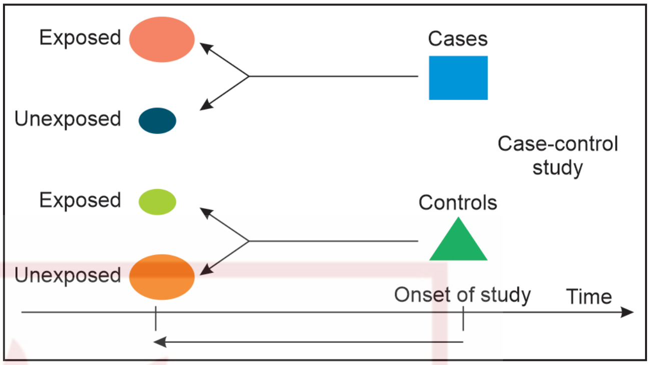


Figure 8.33: **Schematic diagram of case-control study**

Table 8.12 shows a hypothetical scenario of a case-control study.

**Selection of cases:** Cases can be selected from a variety of sources, including hospital patients, clinics, etc. If cases are selected from a single hospital or clinic, risk factors or exposure identified might be unique to that hospital because of referral patterns or other factors thus raising the issue of generalizability. If the study is carried out in a tertiary hospital only, then the risk factors identified might be related to severely ill patients only.

**Selection of controls:** Controls may be selected from the community or from the hospital. In hospitals, they may be selected among patients admitted for diseases other than the disease of interest in the study. In community, controls may be selected from several sources such as voter list, school rosters, resident in the neighborhood sometimes referred to as "neighborhood controls". An alternate method for selecting controls is to use random digit dialing. Another approach for control selection is the "best friend control" in which, the case is asked about the name of a best friend who is likely to participate in the study. Such a control may likely to be similar to the case in age and other demographic and social characteristics. For the studies conducted on genetics, spouse or sibling controls can be used.

**Example:** 100 patients were screened for lung cancer who visited hospital for lung-related disease. 20 patients were found to have developed lung cancer (cases) and 80 patients had no lung cancer (controls). Now, the clinicians collected data regarding the smoking history and the data is placed in the Table 8.13.

Proportion of exposed cases = 15/20 = 0.75

Proportion of exposed controls = 30/80 = 0.38



**Table 8.12** Case control design

| | Cases (with disease) | Controls (without disease) |
|---|---|---|
| Exposed | a | b |
| Not exposed | c | d |
| Proportion of exposed | a/(a + c) | b/(b + d) |

a – Cases who were exposed to risk factor
b – Controls who were exposed to risk factor
c – Cases who were not exposed to risk factor
d – Controls who were not exposed to risk factor

We expect the proportion of cases who were exposed, i.e., a/(a + c), to be greater than the proportion of the controls who were exposed, i.e., b/(b + d)

**Legends used in mixed method research designs:**

| Notation used | Description |
|---|---|
| **QUAN** | Primarily driven by quantitatively data collection |
| **QUAL** | Primarily drive by qualitative data collection |
| **quan** | Quantitative data |
| **qual** | Qualitative data |
| **+** | Indicates that quantitative and qualitative data are collected simultaneously (concurrently) |
| → | Indicates sequential form of data collection |
| **( )** | Indicates that one form of data collection is embedded within another larger design |
| **C** | Quantitative data collection |
| **Q** | Qualitative data collection |
| C+Q | Concurrent data collection |
| Q → C | Sequential data collection, where qualitative data collected first |
| C → Q | Sequential data collection, where quantitative data collected first |

## Types of Mixed Method Research Designs

There are four major types of mixed method research designs:

1. The convergent parallel design
2. The explanatory sequential design
3. The exploratory sequential design
4. The embedded design

### The Convergent Parallel Design

Convergent parallel design is a mixed methods research design that involves the simultaneous collection and analysis of both quantitative and qualitative data (Fig. 8.38). This approach allows researchers to converge or combine data from different sources to gain a more complete understanding of the research topic. Researchers collect quantitative data through surveys, experiments, or other statistical methods, as well as qualitative data through interviews, focus groups, or observations. The data are analyzed separately using appropriate methods for each type of data, and then the results are compared and integrated to draw conclusions about the research question.

The researcher collects both quantitative and qualitative data concurrently. The two types of data are analyzed independently using relevant analytical methods. Subsequently, the outcome of the two analyzes are merged and integrated during the interpretation phase to yield a comprehensive understanding of the research problem. This is further explained using the following nursing research example.

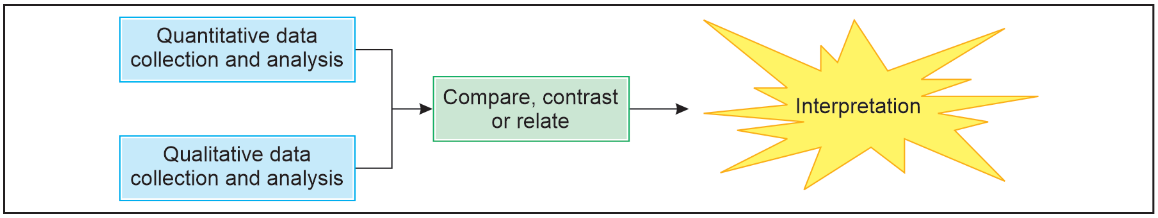


Figure 8.38: **Schematic diagram of convergent parallel design**

**Example**

A nursing researcher is interested in studying the effectiveness of a new pain management intervention for postoperative patients. In this study, the quantitative aspect of the study could involve collecting numerical data on patients' pain levels before and after the intervention using standardized pain assessment tools. This data could be analyzed using statistical methods to determine if the intervention has a significant effect on reducing pain levels. The qualitative aspect of the study could involve conducting interviews with patients to gather their perceptions of the intervention, including their experiences with pain and the effectiveness of the intervention in managing their pain. The qualitative data could be analyzed using thematic analysis to identify common themes and patterns in the patients' experiences.

The results of the quantitative and qualitative analyzes could then be compared and integrated to provide a comprehensive understanding of the effectiveness of the pain management intervention. By using a convergent parallel design, the study could provide more robust evidence on the effectiveness of the intervention by triangulating findings from both quantitative and qualitative data sources.

### The Explanatory Sequential Design

The explanatory sequential design is a research design that involves the collection and analysis of two types of data sets in sequence, with the first phase involving the collection and analysis of quantitative data, followed by a second phase that involves the collection and analysis of qualitative data. This approach is used when the investigator seeks to follow up the quantitative results with qualitative data.

The initial quantitative phase serves to test hypotheses or to explore relationships among variables, and the results are used to develop or refine a theory. The subsequent qualitative phase seeks to further explain or elaborate on the quantitative findings by exploring the experiences, perspectives, or meanings associated with the data.

While the sequence of the approach remains constant, the primary focus of the study may be either quantitative data or qualitative data (or both equally) depending on the research question and objectives (Fig. 8.39).

**Example of an explanatory sequential design in nursing research**

In a study to assess the impact of a new discharge education program on patients with heart failure, the initial quantitative phase involves collecting numerical data to test whether the new discharge education program has a significant effect on clinical outcomes such as readmission rates and medication adherence. The quantitative findings provide a preliminary explanation of the relationship between the intervention and the outcomes. The subsequent qualitative phase involves collecting data on patients' experiences and perspectives to further explain or elaborate on the quantitative findings. The qualitative data may reveal additional factors that contribute to the outcomes, such as patients' understanding of heart failure, their perceived barriers to self-care, and their satisfaction with the education program. The qualitative findings serve to deepen the understanding of the quantitative data and may lead to the development of new theories or hypotheses.



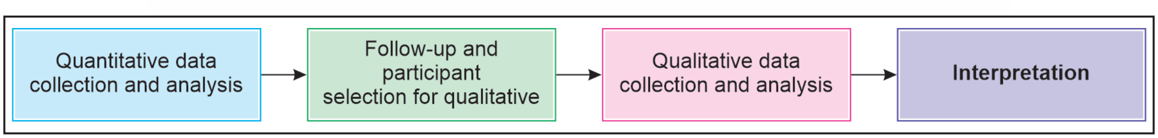


Figure 8.39: **Schematic diagram of explanatory sequential design**

- **Stratification:** When subgroups are present within the study's population, it is necessary to employ stratified sampling to ensure the proper representation of each subgroup in the sample.
- **Consideration for peer review:** It is advisable to undergo peer review or seek consultation from a statistician to validate the appropriateness of sampling methods and to minimize potential errors in the research process.

**By implementing these strategies, you can reduce the risk of sampling errors in nursing research and increase the validity and reliability of your study's findings.**

## SAMPLING TECHNIQUES

In order to obtain the sample, it is necessary to use a systematic procedure for its selection. "**Sampling** is the procedure of selecting a representative subset of a population with the aim of discerning the characteristics of the entire population." Sampling involves the selection of an unbiased or random sample from the population to gain insights into the target population.

The two main types of sampling techniques (Fig. 9.6) are as follows:

1. Probability sampling technique
2. Nonprobability sampling technique

### 1. Probability Sampling Technique

The probability sampling technique ensures that each element, individual, or unit in the population has an equal chance of being selected as a sample. This approach allows for the generation of an unbiased estimate of the population, aiming to closely resemble the characteristics of the entire population. The probability sample is typically chosen randomly from the population. When every element in the population has the same probability of selection, it is known as an "equal probability of selection" (EPS) design. Such designs are also termed "self-weighting" because all sampled units are assigned the same weight.

#### *Advantages of Probability Sampling Technique*

- The probability sampling technique makes valid generalization to the population from which the samples are drawn.
- Probability samples are usually more representative, i.e., they have higher external validity because of less bias.
- More accuracy of sample is there.



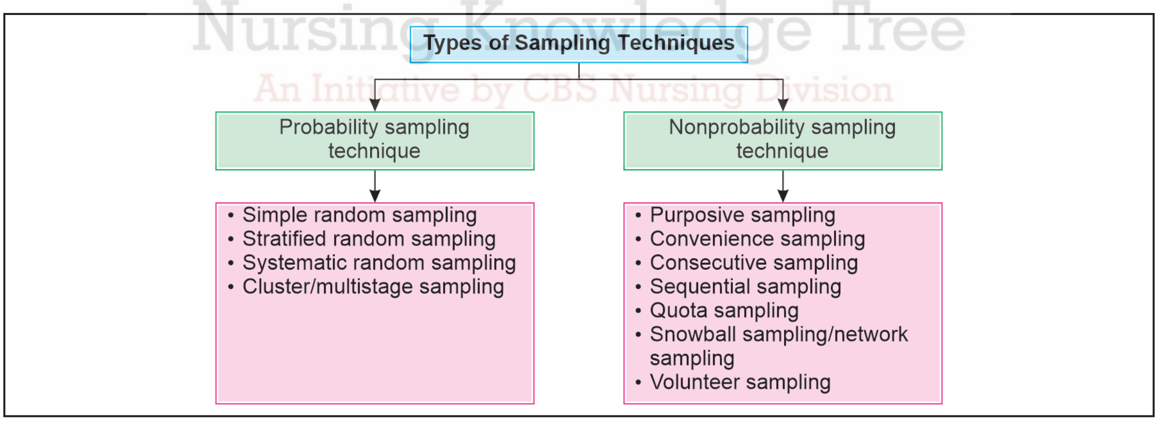


Figure 9.6: **Types of sampling techniques**

## Disadvantages of Probability Sampling Technique

- There is no guarantee of generalization.
- It only allows the researcher to estimate the chance that any given population element will be included in the study.

## Types of Probability Sampling Technique

There are basically four types of probability sampling techniques: (1) Simple random sampling (SRS), (2) Systematic random sampling, (3) Stratified random sampling, and (4) Cluster/Multistage sampling sampling (Fig. 9.7).

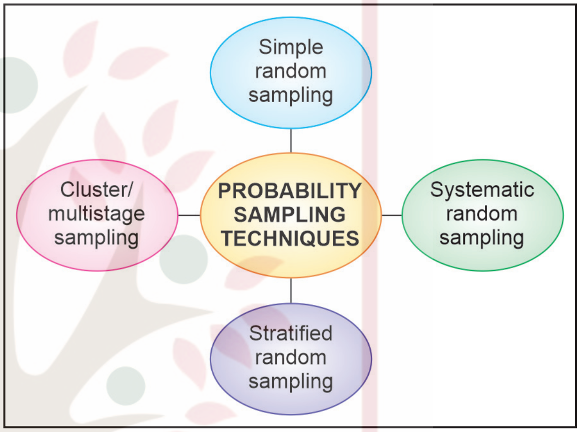


Figure 9.7: **Probability sampling techniques**

### 1. Simple Random Sampling

It is the simplest form of **probability sampling** (Box 9.1). In this method, each individual/subject/unit is chosen randomly and entirely by chance, such that each individual has the **same probability** of being chosen at any stage during the sampling process. Each subset of *k* individuals has the same probability of being chosen for the sample as any other subset of *k* individuals. This **minimizes** bias (Fig. 9.8).

To randomly choose the sample, the researchers can make use of a **lottery method, table of random numbers, computer-generated random number table,** etc. They continue choosing elements at random according to the selected method until the predetermined sample size is attained. Sampling is done without replacement to avoid the occurrence of a unit being sampled more than once.

- **Lottery method:** In the **lottery method**, all items in the sampling frame are numbered on separate slips of paper of identical size and shape. These slips are then folded and mixed up in a bowl. A blindfold selection is then made of the number of slips required to constitute the desired sample size. For example, if we want to take a sample of 10 persons out of a population of 100, then write the names of the 100 persons on separate slips of paper, fold these slips, mix them thoroughly and then make a blindfold selection of 10 slips. In the same way, if we want to know the status of cleanliness of wards in a hospital, all the wards can be assigned a number and their numbered slips can be put in a bowl. The researcher then picks up the slips randomly to choose the wards. Similarly, to know the prevalence of anemia among nurses, they can be selected randomly to estimate their Hb levels.



**BOX 9.1 Steps of simple random sampling technique**

1. Define the target population.
2. Develop a sampling frame.
3. Assign a unique number to each element in the frame.
4. Determine the sample size.
5. Randomly select the targeted number of population elements.

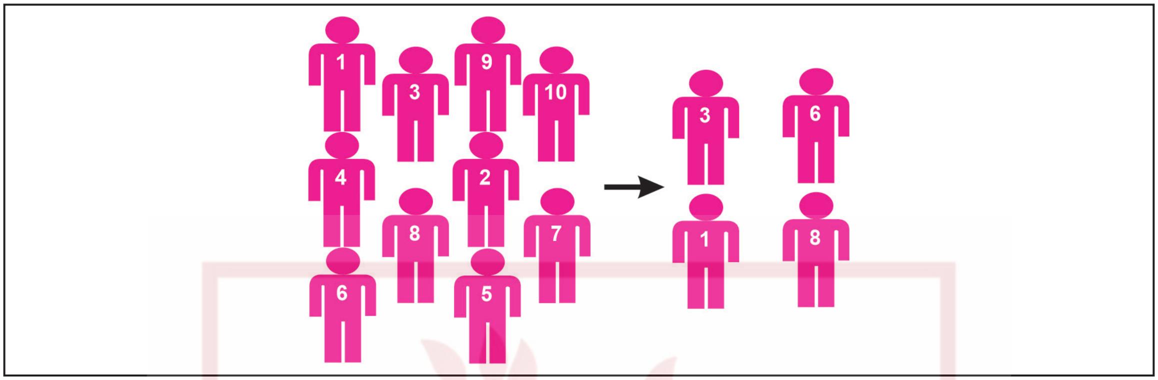


Figure 9.8: **Simple random sampling**

- **Random number table:** A **random number table** consists of 6 columns of 2 or more digit nonrepeatable numbers (Annexure). The researcher, with eyes closed, then points to any number with in the random number table and moves to any systematic fashion (up, down and sideways) to select the correct number of potential participants.
- **Computer-generated random number table:** Figure 9.9A shows the screenshot of Stat Trek's random number generator. The columns we need to fill to generate the random number table are the "minimum" value; "maximum" value (should be more than the required random numbers); "allowing or preventing the duplicate entries" in the generated table; and the "Seed". "The seed is a number that controls whether the Random Number Generator produces a new set of random numbers or repeats a particular sequence of random numbers. If the text box labeled "Seed" is blank, the Random Number Generator will produce a *different* set of random numbers each time a random number table is created. On the other hand, if some number is entered in the "Seed" text box, the Random Number Generator will produce a set of random numbers based on the value of the Seed. Each time a random number table is created, the Random Number Generator will produce the *same* set of random numbers, until the Seed value is changed."

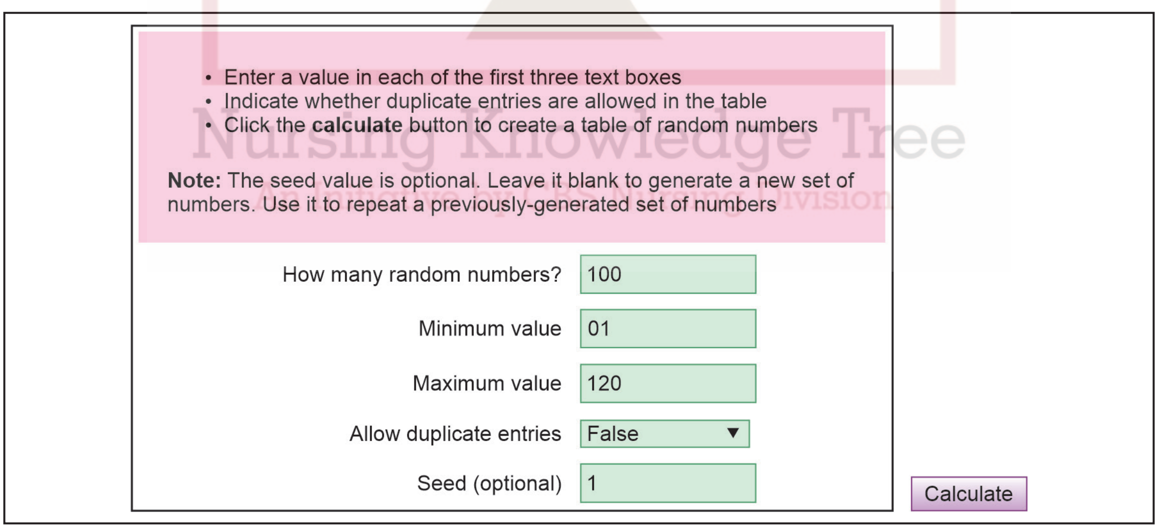




Figure 9.9A: **Screenshot showing Stat Trek's random number generator**

## Types of Diagrams and Graphs

In research studies, various diagrams, charts, and graphs are commonly used to present data in a visually informative and accessible manner. Some of the commonly used diagrams include bar diagrams, Pie charts, Histogram, frequency polygon, line graphs, cumulative frequency curve, Scatter plot, box plot, pictograms and maps.

### 1. Bar Diagram

A bar diagram, also known as a bar chart or bar graph, is a visual representation of data using rectangular bars of varying lengths. It is a widely used method to display and compare data across different categories or groups. Bar diagrams are effective for showing discrete data and making comparisons between data points. Bar diagrams can be either vertical or horizontal. In a *vertical bar diagram,* the numerical values are placed on the y-axis, while the categories or groups are represented on the x-axis. In a *horizontal bar diagram,* the numerical values are placed on the x-axis, and the categories or groups are represented on the y-axis. An example of vertical and horizontal bar diagrams are presented in Figures 14.2 and 14.3 respectively.

There are different type of bar diagrams, each with its specific use and characteristics. The commonly used types are Simple bar diagram, Stacked bar diagram, and Grouped bar diagram (Clustered bar diagram).

- **Simple bar diagram:** A simple bar chart represents data using rectangular bars of uniform width. Each category or group is shown with a single bar, and the height/length of the bar corresponds to the value or quantity for that category. Simple bar charts are useful for comparing data values across different categories. They are particularly effective when discrete or categorical data needs visualization without any stacking or grouping. Bar diagrams presented in Figures 14.2 and 14.3 are example of simple bar diagrams.
- **Stacked bar chart/proportionate bar graph:** Data is displayed as stacked bars, illustrating the composition of each category. Each bar is divided into segments or layers, with each segment representing a different component or subcategory within the main category. The total height of the stacked bar represents the

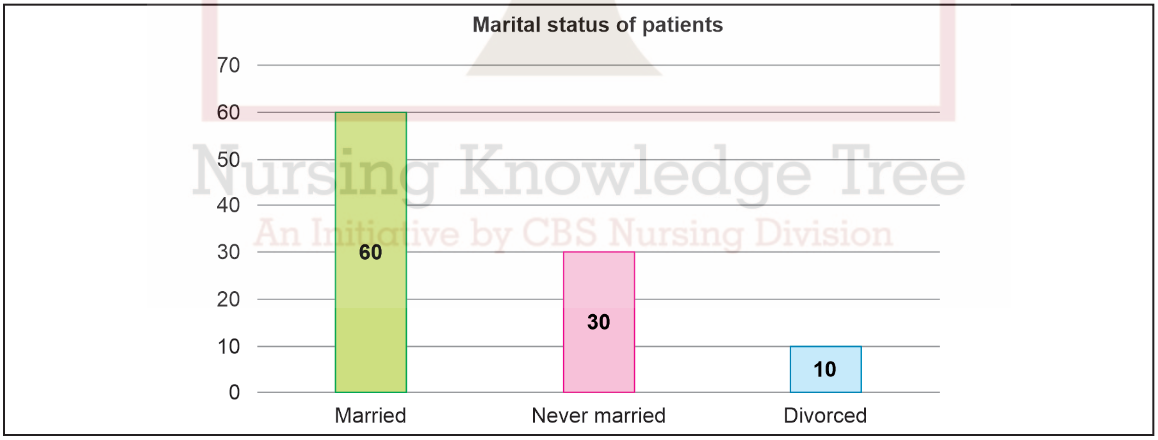


Figure 14.2: **Marital status of patients admitted with COVID-19**

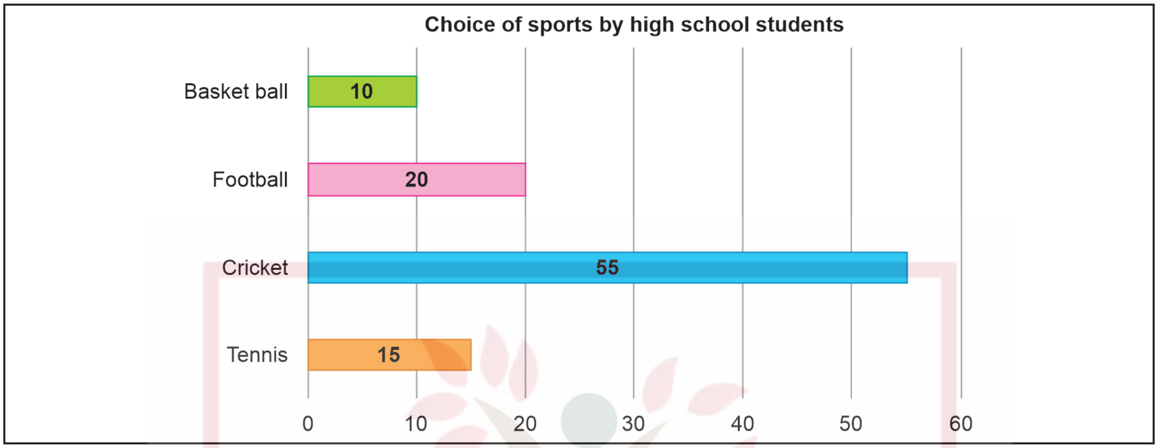


Figure 14.3: **Simple bar diagram showing choice of sports by high school students**

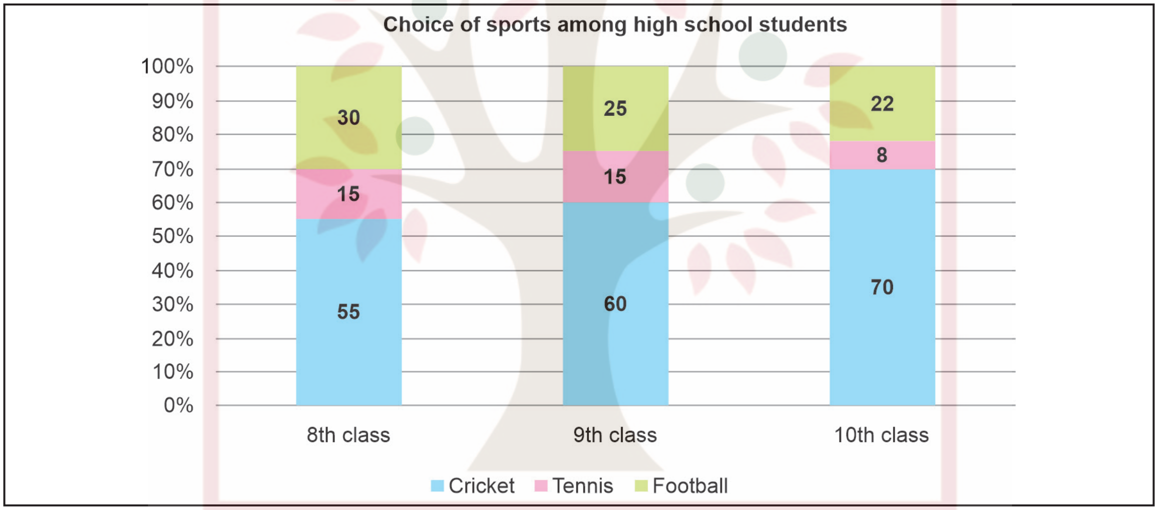


Figure 14.4: **A stacked bar diagram showing choice of sports by 8th, 9th and 10th class high school students**



total value for the category. Stacked bar charts are used for showing part-to-whole relationships and illustrating how different components contribute to the overall total. An example of stacked bar diagram is presented in Figure 14.4.

- **Grouped bar chart/multiple bar chart:** Multiple data series within each category are compared by displaying multiple bars side by side for each category. Each set of bars represents a different data series or subcategory within the main category. Grouped bar charts make it easy to compare the values of each subcategory across the main categories, helping to identify trends and differences. An example of grouped bar chart is presented in Figure 14.5.

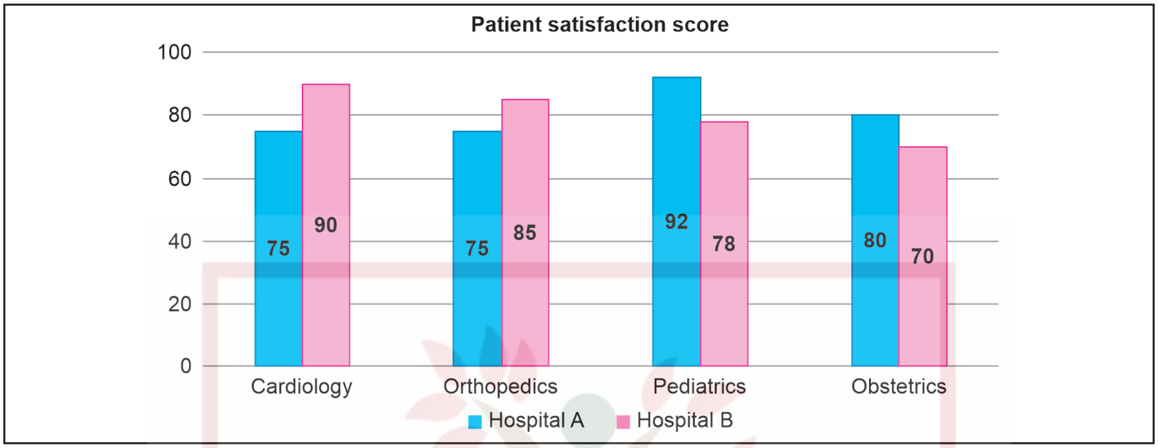


Figure 14.5: **Grouped bar chart (Clustered bar diagram) showing comparison of patient satisfaction in different departments of hospital A and hospital B**

## 2. Pie Chart

A pie chart is a circular chart that is used to represent data in the form of a circle, which is divided into segments, much like slices of a pie. Each segment is represented by a category, and the size of each segment is determined by the quantity or percentage of data it represents. Pie charts are commonly used to display the composition or distribution of a whole in various categories. They are particularly useful when the relative contribution of each category to the whole needs to be shown. In research pie charts can be used to show comparison, visualize percentages, and highlight dominant categories. An example of a pie chart is given in Figure 14.6.

To calculate the degree of each pie segment, the following formula is used:

Degree of Segment = (Percentage of Data in the Category/100) × 360°

**Example:** Calculate the percentage and draw a pie chart for educational status of 60 patients as given in Table 14.6.

**Table 14.6** **Frequency of educational status of patients (N = 60)**

| Educational status of patient | f |
|---|---|
| Illiterate | 15 |
| Primary school | 10 |
| High school | 25 |
| Graduate | 10 |



**Now, let's calculate the percentages for each category:**

Percentage of each category can be calculated by using the following formula:

Percentage of the category = (class frequency/total frequency) × 100

**Percentage of illiterate patients:** (15/60) × 100% = 25%

**Percentage of patients with primary school education:** (10/60) × 100% = 16.67%

**Percentage of patients with high school education:** (25/60) × 100% = 41.67%

**Percentage of graduate patients:** (10/60) × 100% = 16.67%

To draw a pie chart, the degree for each segment of pie must be calculated using the following formula:

Degree of segment = (Class percentage/100) × 360 or Class percentage × 3.6

### Percentage of Illiterate Patients

- Percentage: 25%
- Degree of segment = (25/100) × 360°
- Degree of segment = 90°

### Percentage of Patients with Primary School Education

- Percentage: 16.67%
- Degree of segment = (16.67/100) × 360°
- Degree of segment = 60°

### Percentage of Patients with High School Education

- Percentage: 41.67%
- Degree of segment = (41.67/100) × 360°
- Degree of segment = 150°

### Percentage of Graduate Patients

- Percentage: 16.67%
- Degree of segment = (16.67/100) × 360°
- Degree of segment = 60°

The pie chart shown in Figure 14.6 can be drawn from the above data.

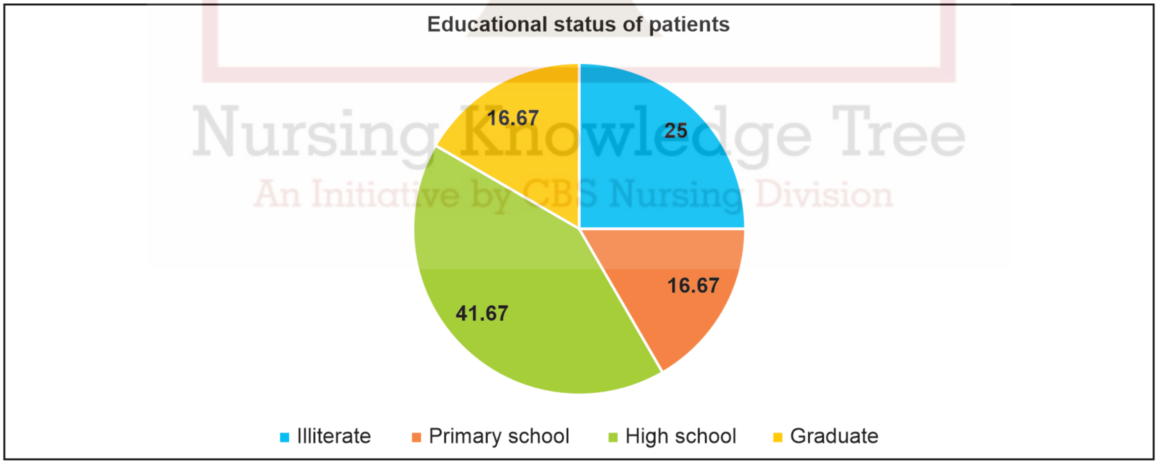


Figure 14.6: **Pie diagram showing educational status of patients**

**Alternative Method (Direct Method)**

Degree for each segment of pie can be directly calculated using the following formula:

Degree of each segment = (Class frequency/total frequency) × 100.

**Applying to the above data:** (total frequency is 60)

- **Degree of Segment for Illiterate Patients:**
  Class frequency: 15
  Degree of segment = (15 / 60) × 360° = 90°
- **Degree of Segment for Patients with Primary School Education:**
  Class frequency: 10
  Degree of segment = (10/60) × 360° = 60°
- **Degree of Segment for Patients with High School Education:**
  Class frequency: 25
  Degree of segment = (25/60) × 360° = 150°
- **Degree of Segment for Graduate Patients:**
  Class frequency: 10
  Degree of segment = (10/60) × 360° = 60°

## 3. Histogram

A histogram is a graphical representation of data that displays the distribution of a dataset. Histograms are best suited for continuous data, where data values fall within a range. For example, they are commonly used to display the distribution of ages, test scores, income levels, or any data that can be measured along a continuous scale.

Histogram consists of a series of adjacent rectangles, or bars, where the width of each bar corresponds to a particular range or interval of data values, and the height of each bar represents the frequency or count of data points within that range. Unlike bar charts, there are no gaps between the bars in a histogram. The bars are adjacent because they represent continuous data intervals. The x-axis (horizontal axis) of a histogram represents the data intervals or bins, while the y-axis (vertical axis) represents the frequency or count of data points in each interval.

**Steps to Draw a Histogram**

A histogram can be drawn using the following steps:

1. **Select data:** Ensure the data is quantitative and continuous, such as ages, test scores, or weight of patient, etc.
2. **Data binning:** Divide the range of data values into intervals or bins. These bins should be mutually exclusive (data points can belong to only one bin) and collectively exhaustive (covering the entire range of data values). The choice of bin size or width is important and can affect the interpretation of the histogram.
3. **Frequency count:** Count the number of data points that fall into each bin. This is done by tallying how many data points have values within each bin's range.
4. **Scale the y-axis:** On the y-axis (vertical axis), represent the frequency or count of data points. The scale should be appropriate for the data, and it is common to use whole numbers (e.g., 10, 20, 30) depending on the data and the number of bins.

5. **Scale the x-axis:** On the x-axis (horizontal axis), represent the bins or intervals. Each bin is usually labeled with the starting and ending values of the range it covers.
6. **Bar construction:** For each bin, draw a rectangular bar that starts at the lower boundary of the bin and extends to the upper boundary. The width of each bar should correspond to the width of the bin, and the height of the bar represents the frequency count of data points within that bin.
7. **Labeling:** Label the axes with appropriate titles and include units if applicable. This helps readers understand what the histogram represents.
8. **Title:** Include a title for the histogram that briefly describes the dataset and the variable being represented.

**Example:** Draw a histogram for the following grouped data set

| Height in cm | No. of patients |
|---|---|
| 140–145 | 6 |
| 145–150 | 12 |
| 150–155 | 18 |
| 155–160 | 25 |
| 160–165 | 32 |
| 165–170 | 30 |
| 170–175 | 23 |
| 175–180 | 15 |
| 180–185 | 10 |
| 185–190 | 5 |

In the given example, the data has been already grouped into bins (intervals). Histogram can be prepared from the data using the steps described from 4 to 8 as described above. The histogram is presented in Figure 14.7.

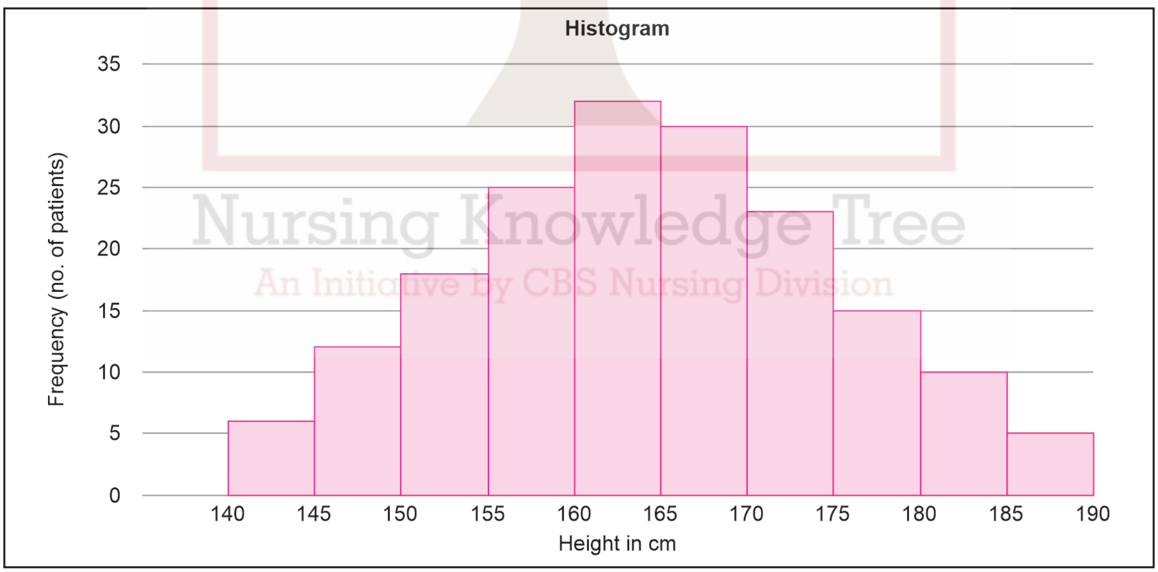


Figure 14.7: **Histogram depicting height (in cm) of patients**

## 4. Frequency Polygon

A frequency polygon is a graphical representation of data that displays the distribution of a dataset using lines or curves. A frequency polygon can be obtained by connecting the midpoints of top of the bars in a histogram. This is a common method for creating a frequency polygon from a histogram, and it provides a way to represent the same data in a different visual format. In some cases, the line connecting the points is smoothed to create a curve, especially when the data points are grouped into narrower intervals. The curve provides a more continuous and visually appealing representation.

When creating a frequency polygon from a histogram, it is common practice to connect the end points of the frequency polygon lines to the horizontal axis (x-axis). This ensures that the polygon is a closed figure and that it connects back to the x-axis. An example of frequency polygon is presented in Figure 14.8.

**A frequency polygon can be drawn from histogram using the following steps:**

1. **Draw histogram:** Construct the histogram as usual.
2. **Midpoint calculation:** Calculate the midpoint for each bin, typically represented as the average of the lower and upper boundaries of the bin.
3. **Point plotting:** The midpoints are plotted on the x-axis, with each midpoint serving as the center of a bin in the histogram.
4. **Frequency count:** The corresponding heights (y-values) for each midpoint on the frequency polygon are determined using the frequency counts from the histogram (midpoint at top of the bin).
5. **Connection of points:** The midpoints are connected by lines or curves. The initiation of the connection is made by linking the first midpoint in x-axis. This action results in the representation of the frequency polygon.
6. **Closing of the polygon:** To finalize the frequency polygon, the last point is connected back to the horizontal axis (x-axis). This step serves to close the shape, creating a complete representation.
7. **Labeling and titling:** The x-axis and y-axis are labeled appropriately, including the inclusion of units of measurement. Additionally, a title is provided for the frequency polygon.

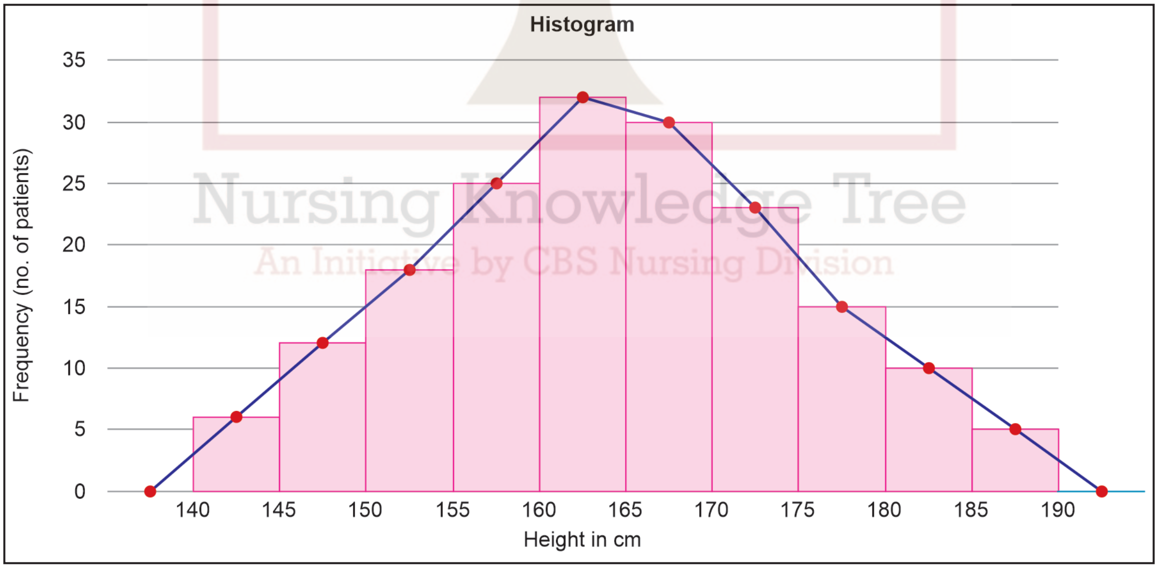


Figure 14.8: **Frequency polygon depicting heights (in cm) of patients**

## 5. Line Graphs

A line graph, also known as a line chart, is a graphical representation of data that uses lines to connect individual data points. It is commonly used to show trends and changes in data over time or to illustrate the relationship between two variables. Line graphs are often used to display time series data, where data points are collected at regular intervals over time. This can include data on temperature measurements, population growth, inflation rate and more.

Line graphs typically represent the relationship between two variables. One variable is plotted on the x-axis (horizontal), and the other variable is plotted on the y-axis (vertical). For example, time may be on the x-axis, and the variable being measured over time is on the y-axis. Individual data points are marked on the graph, and lines connect these points in chronological order. The lines represent the trend or pattern in the data. In a line graph, more than one variable can be compared by plotting multiple lines on the same graph. This allows for the visual analysis and comparison of the trends and relationships between multiple variables over the same time period or along the same x-axis. An example line graph is presented in Figure 14.9.

| Month/2022 | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Delhi (°C) | 14 | 17 | 23 | 30 | 34 | 33 | 32 | 31 | 31 | 28 | 22 | 18 |
| Chennai (°C) | 24 | 25 | 28 | 32 | 35 | 34 | 32 | 31 | 31 | 29 | 27 | 25 |

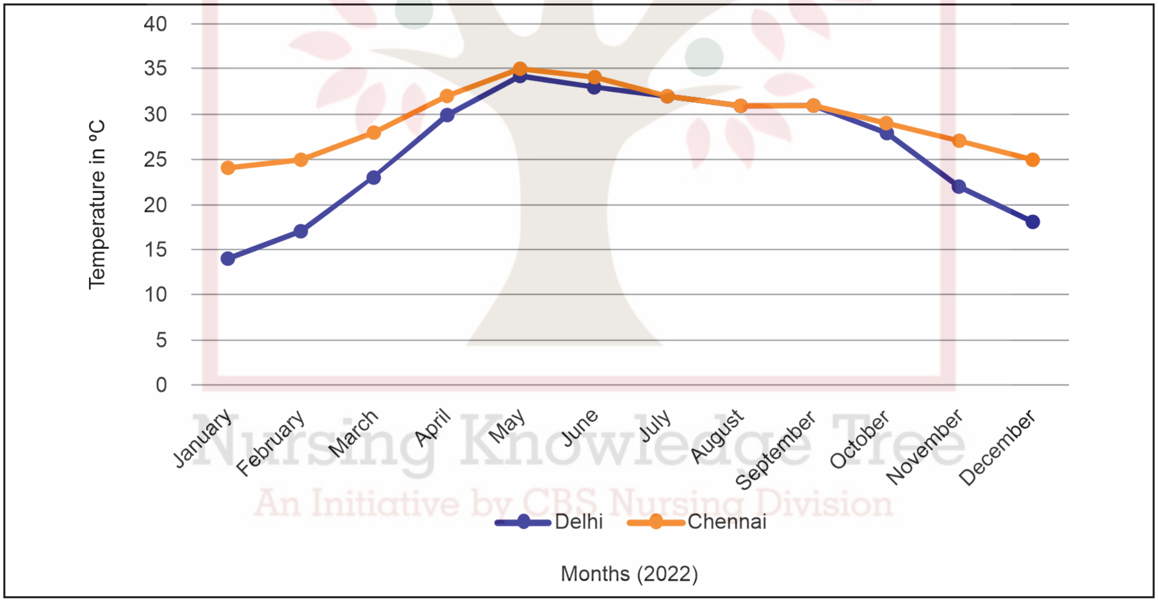


Figure 14.9: **Line graph depicting average temperature of Delhi and Chennai from January 2022 to December 2022**



STATISTICS

CHAPTER 15

# Descriptive Statistics: Measures of Central Tendencies and Variability

## LEARNING OBJECTIVES

After the completion of the chapter, the readers will be able to:
- Interpret the measures of central tendencies and their purpose.
- Compute the mean, median and mode of grouped and ungrouped data.
- Explain the measures of dispersion and their purpose.
- Compute the measures of dispersion of grouped and ungrouped data.

## CHAPTER OUTLINE

- Descriptive Statistics
- Measures of Central Tendency
- Measures of Variability/Dispersion

## KEY TERMS

**Arithmetic mean:** The arithmetic mean is the sum of a set of values divided by the number of values in the set, representing the typical value in the dataset.

**Descriptive statistics:** Descriptive statistics is a branch of statistics that focuses on summarizing and presenting data in a meaningful and interpretable way.

**Interquartile range:** The interquartile range is a measure of statistical dispersion, indicating the spread between the first and third quartiles of a dataset.

**Mean:** The mean is the average value of a dataset, calculated by summing all values and dividing by the number of values.

**Measure of central tendency:** A measure of central tendency is a statistical value that represents the center or typical value in a dataset, with examples including the mean, median, and mode.

**Measure of dispersion:** A measure of dispersion is a statistical value that quantifies the spread or variability of data in a dataset, with examples including the range, variance, and standard deviation.

*Contd...*

**Median:** The median is the middle value in a dataset when values are ordered (ascending or descending), separating the higher half from the lower half.

**Mode:** The mode is the most frequently occurring value in a dataset, representing the value that appears most often.

**Quartiles:** Quartiles are values that divide a dataset into four equal parts, with the first quartile ($Q_1$) at the 25th percentile, the second quartile ($Q_2$) as the median (50th percentile), and the third quartile ($Q_3$) at the 75th percentile.

**Range:** The range is the difference between the maximum and minimum values in a dataset, indicating the extent of variation.

**Standard deviation:** The standard deviation is a measure of the amount of variation or dispersion in a dataset, providing insights into the spread of data values.

**Variance:** Variance is a measure of how much data points deviate from the mean of a dataset, calculated as the average of the squared differences from the mean.

## DESCRIPTIVE STATISTICS

Descriptive statistics is a branch of statistics that focuses on summarizing and presenting data in a meaningful and interpretable way. It involves using various techniques and measures to describe, organize, and visualize data, enabling researchers, analysts, and decision-makers to gain insights and make informed conclusions.

Descriptive statistics serves as the first step in data analysis. It helps condense large datasets into a manageable form while retaining the essential information. Descriptive statistics often involves data visualization techniques, such as histograms, box plots, and bar charts, to provide a graphical representation of data distributions and patterns.

Descriptive statistics focuses on summarizing data. It serves as a foundation for inferential statistics. Summarization techniques include:

- Measures of central tendency
- Measures of variability/dispersion.

## MEASURES OF CENTRAL TENDENCY

In research studies, we gather large amount of data with many observations. A single number representative of all these observations needs to be identified. Such a number can be a central value or an "average" of the data.

Measures of central tendency are statistical measures that provide insight into the central or typical value in a set of data. The objective of an average is to provide a single central value which would be representative of the entire mass of observations. It makes the comparison easier. They help researchers and analysts understand where the data tends to cluster or center. The three most common measures of central tendency are:

i. **Mean:** The mean, often referred to as the average, is calculated by summing up all values in a dataset and dividing by the total number of values. It represents the arithmetic center of the data.

ii. **Median:** The median is the middle value in a dataset when the values are ordered from smallest to largest or vice versa. If there is an even number of value, the median is the average of the two middle values. The median represents the value that divides the data into two equal halves.

iii. **Mode:** The mode is the value that appears most frequently in a dataset. In some cases, there may be more than one mode, making the dataset multimodal.

There are three main types of averages which include:

1. **Mathematical average:** Average calculated using mathematical equations. It includes arithmetic mean, geometric mean and harmonic mean.
2. **Positional average:** Its value is calculated by locating its position in the series. It includes Median, Mode, Quartiles, and Percentiles, etc.
3. **Miscellaneous average:** It includes moving average and progressive average.

## Arithmetic Mean

It is the most commonly used method for obtaining "average". It is defined as sum total of all the observations divided by the number of observations.

### *Formula for Mean/Arithmetic Mean*

$$\bar{X} = \frac{\Sigma X}{N} \qquad \ldots(1)$$

$$\bar{X} = (\Sigma X)/N$$

$$\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{N}$$

In this formula, $X$ = Arithmetic Means
$\sum X$ = Sum of all observations, i.e., $X_1 + X_2 + \ldots \ldots + X_n$
$N$ is the total number of observations.

## Arithmetic Means for Individual Observations

First add up all the observations of given data and then, divide by the total number of observations as in the following illustration:

**Illustration 1:**

Calculate mean of the given data: 10, 12, 11, 12, 13, 14, 15, 14, 14, 16, 14, 17, 18, 9, 15

**Solution:**

Here $N = 15$

$\Sigma X = 10 + 12 + 11 + 12 + 13 + 14 + 15 + 14 + 14 + 16 + 14 + 17 + 18 + 9 + 15 = 204$

Putting the formula,

$$\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{N}\text{, we get}$$

$$\bar{X} = \frac{204}{15}$$

$$= 13.6$$

So, mean of given data is 13.6.

**Arithmetic Mean in Case of Discrete Series**

$$\bar{X} = \frac{\Sigma f X}{\Sigma f} \quad \ldots(2)$$

*Where,* $\bar{X}$ = *Arithmetic mean*

$\Sigma fX$ = *Sum total of variable multiplied by corresponding frequencies, i.e.,* $fX_1 + fX_2 + \ldots + fX_n$

$\Sigma f$ *is the total number of observations.*

**Method**

- Multiply the frequency ($f$) of each row by variable ($X$) to get $fX$
- Adding all these will give $\Sigma fX$
- Divide $\Sigma fX$ by $\Sigma f$ to get mean.

**Frequency** is the number of times a value occurs in a data set. In the above illustration, score of 12, 14, and 15 have appeared 2 times, 4 times and 2 times, respectively.

Instead of adding 12 two times; 14 four times and 15 two times, we facilitate our calculation by the multiplication of 12 by 2, 14 by 4 and 15 by 2 and obtain the Mean ($\bar{X}$) as follows:

$$\bar{X} = \frac{[10 + (12 \times 2) + 11 + 13 + (14 \times 4) + (15 \times 2) + 16 + 17 + 18 + 9]}{15}$$

$$\bar{X} = \frac{10 + (12 \times 2) + 11 + 13 + (14 \times 4) + (15 \times 2) + 16 + 17 + 18 + 9]}{15}$$

$$\bar{X} = \frac{[10 + (12 \times 2) + 11 + 13 + (14 \times 4) + (15 \times 2) + 16 + 17 + 18 + 9]}{15}$$

$$\bar{X} = \frac{204}{15} = 13.6$$

This method is little tedious. It is better to construct a frequency distribution table first, as is shown in the following illustration.

**Illustration 2:**

Calculate the mean of the following data table:

| Marks ($X$) | 3 | 4 | 5 | 6 | 7 | 8 | 10 | 11 | 13 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|
| No. of Students ($f$) | 2 | 4 | 3 | 8 | 9 | 7 | 7 | 6 | 4 | 3 |



In above data table, marks obtained by the students act as variable ($X$) and the number of students who obtain a particular value of ($X$) is the corresponding frequency ($f$).

**Now, to calculate Arithmetic Mean in this case we need.**

- $\Sigma fX$, i.e., sum of product of $f$ and $X$
- $\Sigma f$, i.e., the sum of all frequencies

| Sl. no. | Marks ($X$) | No. of students ($f$) | $fX$ |
|---|---|---|---|
| 1 | 3 | 2 | 6 |
| 2 | 4 | 4 | 16 |
| 3 | 5 | 3 | 15 |
| 4 | 6 | 8 | 48 |
| 5 | 7 | 9 | 63 |
| 6 | 8 | 7 | 56 |
| 7 | 10 | 7 | 70 |
| 8 | 11 | 6 | 66 |
| 9 | 13 | 4 | 52 |
| 10 | 15 | 3 | 45 |
| | | $\Sigma f = 53$ | $\Sigma fX = 437$ |

Here $\Sigma fX = 437$ and $\Sigma f = 53$

So, $\bar{X} = \dfrac{437}{53} = 8.24$

**Arithmetic Mean for Continuous Series**

In this case, mid points of class intervals are taken as "$X$". These are multiplied by their corresponding frequencies "$f$" to get $fX$. These are summed up to get $\Sigma fX$. Then, use the following formula to calculate mean.

$$\bar{X} = \frac{\Sigma fX}{\Sigma f}$$

**Illustration 3:**

Calculate mean of marks obtained by students. Use data from table given below:

| Marks ($X$) | No. of students ($f$) |
|---|---|
| 10–20 | 5 |
| 20–30 | 7 |
| 30–40 | 11 |
| 40–50 | 8 |



**Solution:**

- First, get mid point for each class limit using the formula

$$\frac{\text{Upper class limit + Lower class limit}}{2}$$

**For example:** Mid-point for the class interval 10–20:

$$\frac{10+20}{2} = 15$$

Similarly, calculating for each class interval, we get the following table:

| Marks | Mid-point ($X$) | No. of students ($f$) |
|---|---|---|
| 10–20 | 15 | 5 |
| 20–30 | 25 | 7 |
| 30–40 | 35 | 11 |
| 40–50 | 45 | 8 |

- Then calculate ($fX$) for each row, i.e., 15 × 5 = 75, 25 × 7 = 175, 35 × 11 = 385, and 45 × 8 = 360
- Calculate $\Sigma f$
- Calculate $\Sigma fX$

| Marks | Mid-point ($X$) | No. of students ($f$) | $fX$ |
|---|---|---|---|
| 10–20 | 15 | 5 | 75 |
| 20–30 | 25 | 7 | 175 |
| 30–40 | 35 | 11 | 385 |
| 40–50 | 45 | 8 | 360 |
| **Total** | | $\Sigma f = 31$ | $\Sigma fX = 995$ |

Now, calculate mean using the formula (2).

$$\bar{X} = \frac{\Sigma fX}{\Sigma f}$$

$$\frac{995}{31} = 32.10$$

Hence, the mean is 32.10.

**Merits of arithmetic mean:**

1. It is rigidly defined, therefore chances of its ambiguity are not there.
2. It is easy to calculate.
3. It is based on all the observations in the series.
4. Values in the series need not to be arranged in any order.
5. It is a reliable estimate of a central value as it does not get affected much when repeated samples are taken from same population.
6. It is subjected to further algebraic treatment.
7. It acts as center of gravity, hence balance both the sides in a series.
8. It is based on mathematical calculation and not on any position in the series.

**Demerits of arithmetic mean**

1. It is affected if there are extreme values in the data series.
2. Sometimes it gives a value which is not included in a series.
3. Sometimes it gives a fractional value which is not possible in the concerned case.
   For example, average number of children in 4 families with children 2, 3, 3, 3 respectively would be 2.75 which is impossible.
4. It cannot be calculated in case of a distribution with open ended classes (class having either no upper class limit or lower class limit).
5. It only gives a representative value in case of "bell shaped" curve but not in case of "U" shaped curve.

# Median

It is value of the middle item or observation of a series when the data is arranged in some order of magnitude, i.e., in ascending or descending order. It indicates the value of middle item in the distribution.

It is a positional average and divides the series into two equal parts. Therefore, it is also the value of 50th percentile.

## Median for Individual Observations

**Illustration 4:**

Calculate median for given date set: 7, 6, 14, 17, 1, 5, 7, 9, 3.

**Solution:**

- Arrange the data in ascending or descending order first, i.e., 1, 3, 5, 6, 7, 7, 9, 14, 17.
- If the number of observations is odd then median is the value of the $\frac{n+1}{2}th$ item in the series.
- If the number of $\left(\frac{n+1}{2}th\right)$ observation is even, then we have to take Mean of $\frac{N}{2}th$ item and $\frac{n+1}{2}th$ item. That is,

$$\frac{\frac{N}{2}th \text{ item} + \left(\frac{N}{2}+2\right)th \text{ item}}{2}$$

- In the above example, the number is odd, i.e., $N = 9$ then Median is $\frac{n+1}{2}th$ item

$$= \frac{(9+1)}{2} th \text{ item, i.e., } \frac{10}{2} th = 5\text{th item.}$$

In our data set 5th item is 7. So, median is 7 in given data set.

**Illustration 5:**

Calculate median for given data set: 7, 6, 14, 17, 1, 5, 7, 9

**Solution:**

- Arrange the data in ascending or descending order, i.e., 17, 14, 9, 8, 7, 6, 5, 1
- In this case, number is even, i.e., $N = 8$ then

$$\text{Median} = \frac{\frac{N}{2}th \text{ item} + \left(\frac{N}{2}+1\right)th \text{ item}}{2}$$

$$= \frac{\frac{8}{2}th \text{ item} + \left(\frac{8}{2}+1\right)th \text{ item}}{2} = \frac{4th \text{ item} + (4+1)th \text{ item}}{2}$$

$$= \frac{4th \text{ item} + 5th \text{ item}}{2}$$

- 5th item is in given data set is 7 and 4th item is 8. Putting these values in above formula, we get

$$\frac{8+7}{2}=\frac{15}{2}=7.5$$

So, median is 7.5 for above data.

## Median for Discrete Series

**Illustration 6:**

Calculate median for given data set.

| Marks (*X*) | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|
| No. of students (*f*) | 3 | 5 | 10 | 8 | 6 |

**Solution:**

First calculate cumulative frequency *(cf)* for the above data.

| Marks (*X*) | No. of students (*f*) | Cumulative frequency (*cf*) |
|---|---|---|
| 3 | 3 | 3 |
| 4 | 5 | 5 + 3 = 8 |
| 5 | 10 | 10 + 8 = 18 |
| 6 | 8 | 8 + 18 = 26 |
| 7 | 6 | 6 + 26 = 32 |
| | $\Sigma f = 32$ | |

Since $N=\Sigma f=32$ is even, the Median is $\dfrac{\dfrac{N}{2}th\text{ item} + \left(\dfrac{N}{2}+1\right)th\text{ item}}{2}$

Here, $\dfrac{N}{2}th\text{ item}=\dfrac{32}{2}th\text{ item}=16th\text{ item}$

and, $\left(\dfrac{N}{2}+1\right)th\text{ item}=\left(\dfrac{32}{2}+1\right)th\text{ item}=17th\text{ item}$

Hence, 16th and 17th items come against the cumulative frequency 18.
Its value is 5 in marks column. So, the median for the given data is 5.



## Median for Continuous Series

**Illustration 7:**

Calculate median for given observation:

| Marks | No. of students (*f*) |
|---|---|
| 10–20 | 5 |
| 20–30 | 7 |
| 30–40 | 10 |
| 40–50 | 8 |

# Nursing Research and Statistics

## Salient Features

- Written in adherence with the revised INC syllabus for BSc Nursing also useful for postgraduate students and PhD scholars.
- A practical approach has been adopted to enrich the learning experience with appropriate examples of published research works from reputed journals.
- Supportive Illustrations in the form of images, flowcharts and tables have been provided to simplify the concepts for the students.
- Recent updates and data covered extensively.
- **15+** Infographic Videos included as a Bonus Content.

**Learning Objectives** enlist what the students will learn after studying the entire chapter.

**LEARNING OBJECTIVES**

After the completion of the chapter, the readers will be able to:
- Define research.
- Describe scientific methods of acquiring knowledge.

**Chapter Outline** provides a quick glance of the entire chapter in one go.

**CHAPTER OUTLINE**
- Introduction
- Definitions
- Identification of Research Problem
- Formulating the Research Problem
- Elements of Problem Statement

Enriched with numerous **research examples** from applicability point of view.

**Examples of conceptual and operational definitions of 'nursing competence'**

**Conceptual definition:** Nursing competence represents the comprehensive knowledge, skills, judgment, and ethical principles that underpin a nurse's ability to provide safe, effective and compassionate patient care.

Boxes included in the chapters enrich students with extra information and provide them an edge over others.

**BOX 7.1 Theory versus model**

Theories are general and abstract ideas that explain how things work. Models are specific and concrete ways of showing how things are or should be. Models can come from theories, but they can also come from observations or experiences.

**Illustrations and Tables** are used to make learning easy for students.

STEP 1 Selecting a topic for review
STEP 2 • Deciding the key words • Searching the literature (electronic/manual)
STEP 3 Critiquing the literature by reading and analyzing the literature

**Table 13.2** References versus bibliography

| Reference | Bibliography |
|---|---|
| References are given to the content which is quoted in the article/text or assignment | The bibliography includes list of all the sources which a researcher has gone through to conceive the idea |
| Mostly from the primary data sources | From both primary and secondary data sources |

**ASSESS YOURSELF**

**Long Answer Questions**

2. Discuss five major sources of literature.
3. Discuss five major online sources of literature.

**Short Answer Questions**

1. Define concepts.
2. Define framework.

**Multiple Choice Questions**

1. A blueprint for the conduction of study is:
   a. Pilot study
   b. Research hypothesis
   c. Research design
   d. Data collection tool

Detailed **Assess Yourself** exercises in each and every chapter which will facilitate structured learning and revision of the material provided in the respective chapters.

**Bibliography** allows students to learn more on nursing research and enhance their efficiency.

**BIBLIOGRAPHY**

- *Abbott P. Why do we Need to Review Literature? Nurse Researcher 1993; 1(1): 14-22.*
- *Burns N, Grove S. The Practice of Nursing Research: Conduct, Critique, and Utilization.(4th ed.), 2001, Philadelphia: W.B. Saunders Company.*

## About the Authors

**Sukhpal Kaur** PhD, MSc (N), FNRSI is Principal at National Institute of Nursing Education (NINE), PGIMER, Chandigarh, India. She has a considerable teaching experience and a number of awards and honors to her credit. She is member of editorial board of various national and international journals. She has published about two hundred articles in reputed journals.

**Muthuvenkatachalam Srinivasan** PhD (INC Consortium), MSc Nursing (AIIMS, New Delhi), D Pharm, RN (NMBA, Australia) is currently serving as an Associate Professor at the College of Nursing, All India Institute of Medical Sciences (AIIMS), Mangalagiri, Andhra Pradesh. Previously, he held the position of Assistant Professor at the Institute of Human Behaviour and Allied Sciences (IHBAS), New Delhi. Earlier, he had served as an Associate Professor at the College of Nursing, AIIMS, New Delhi.

CBSPD
Dedicated to Education

**CBS Publishers & Distributors Pvt. Ltd.**
4819/XI, Prahlad Street, 24 Ansari Road, Daryaganj, New Delhi 110 002, India
**E-mail:** feedback@cbspd.com, **Website:** www.cbspd.com
New Delhi | Bengaluru | Chennai | Kochi | Kolkata | Lucknow | Mumbai
Hyderabad | Jharkhand | Nagpur | Patna | Pune | Uttarakhand

ISBN: 978-93-90619-91-7
9 789390 619917