

PEMILIHAN UJI STATISTIK KEDOKTERAN

A full-page background image of a male scientist in a white lab coat, wearing a white hairnet and safety goggles. He is looking through a large, black, professional-grade microscope. His hands, wearing white gloves, are holding a test tube and pouring its contents into a larger Erlenmeyer flask. In the foreground, there are several other glassware items, including a beaker and another flask, on a lab bench. The entire scene is bathed in a cool blue light, with soft bokeh effects in the background.

Prof. Dr. Drs. H. Kusmardi, M.Sc., CIRR.
Prof. Dr. Drs. Yurnadi, M.Kes.
Dr. Andi Muh. Maulana, M.Sc., PAK.

 **Pemilihan Uji Statistik**

Kedokteran



Prof. Dr. Drs. H. Kusmardi, M.Sc., CIRR.

Prof. Dr. Drs. Yurnadi, M.Kes.

Dr. Andi Muh. Maulana, M.Sc., PAK.

Dilarang diperjualbelikan naskah
hasil layanan ini. Dilarang menerbit





www.penerbitbukumurah.com

**Dilarang keras, mencetak naskah
hasil layout ini tanpa seijin Penerbit**

PENERBIT KBM INDONESIA

adalah penerbit dengan misi memudahkan proses penerbitan buku-buku penulis di tanah air Indonesia, serta menjadi media *sharing* proses penerbitan buku.

PEMILIHAN UJI STATISTIK KEDOKTERAN

Copyright @ 2026 By Prof. Dr. Drs. H. Kusmardi, M.Si., CIRR, dkk

All right reserved

Penulis

Prof. Dr. Drs. H. Kusmardi, M.Sc., CIRR.

Prof. Dr. Drs. Yurnadi, M.Kes.

Dr. Andi Muh. Maulana, M.Sc., PAK.

Desain Sampul

Eca Art

Tata Letak

Husnud Diniyah

Editor

Dr. Muhamad Husein Maruahey, Drs., MSc.

Background isi buku di ambil dari [https://www. freepik. com/](https://www.freepik.com/)

Official

Depok, Sleman-Jogjakarta (Kantor)

Penerbit KBM Indonesia

Anggota IKAPI/No. IKAPI 279/JTI/2021

081357517526 (Tlpn/WA)

www.penerbitbukumurah.com

Website

[https://penerbitkbm. com](https://penerbitkbm.com)

[www. penerbitbukumurah. com](http://www.penerbitbukumurah.com)

Email

[naskah@penerbitkbm. com](mailto:naskah@penerbitkbm.com)

Distributor

[https://penerbitkbm. com/toko-buku/](https://penerbitkbm.com/toko-buku/)

Youtube

Penerbit KBM Sastrabook

Instagram

@penerbit. kbmindonesia

@penerbitbukujogja

ISBN: 978-634-278-184-5

Cetakan ke-1, Februari 2026

15,5 x 23 cm, viii + 364 halaman

Isi buku diluar tanggungjawab penerbit
Hak cipta merek KBM Indonesia sudah terdaftar di DJKI-Kemenkumham
dan isi buku dilindungi undang-undang.

Dilarang keras menerjemahkan, memfotokopi, atau
memperbanyak sebagian atau seluruh isi buku ini
tanpa seizin penerbit karena beresiko sengketa hukum
www.penerbitbukumurah.com

**Dilarang keras, mencetak naskah
hasil layout ini tanpa seizin Penerbit**



www.sanksipelanggaranpasal113.com

Undang-Undang No. 28 Tahun 2014 Tentang Hak Cipta

- i. Setiap Orang yang dengan tanpa hak melakukan pelanggaran hak ekonomi sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf i untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 1 (satu) tahun dan/atau pidana denda paling banyak Rp 100. 000. 000 (seratus juta rupiah).
- ii. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp 500. 000. 000,00 (lima ratus juta rupiah).
- iii. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf a, huruf b, huruf e, dan/atau huruf g untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 4 (empat) tahun dan/atau pidana denda paling banyak Rp 1. 000. 000. 000,00 (satu miliar rupiah).
- iv. Setiap Orang yang memenuhi unsur sebagaimana dimaksud pada ayat (3) yang dilakukan dalam bentuk pembajakan, dipidana dengan pidana penjara paling lama 10 (sepuluh) tahun dan/atau pidana denda paling banyak Rp 4. 000. 000. 000,00 (empat miliar rupiah).

Halaman Persembahan

Buku ini dipersembahkan untuk cucuku: Kira Myiesha Alfaliza, Yang baru saja berhasil meraih Juara 2 Olimpiade Sains Puskantara 2026 tingkat nasional dan Juara 2 Olimpiade Matematika POSBANAS tingkat nasional. Semoga kesukaan pada sains selalu dipupuk dengan daya juang tangguh untuk bisa istiqomah memperdalam keilmuan. Semoga dalam jangka pendek semangatmu bisa masuk ke jenjang pendidikan tinggi yang kau impikan dan banggakan, dan jangka panjang minatmu kelak bisa mengantarkan insan manusia menjadi ke generasi semakin maju yang tentunya juga bisa menyeimbangkan antara keilmuan dan ketaqwaan



www.penerbitbukumurah.com

**Dilarang keras, mencetak naskah
hasil layout ini tanpa seijin Penerbit**



PRAKATA

Buku ajar ini disusun sebagai respons terhadap tantangan fundamental yang dihadapi oleh para peneliti, klinisi, dan mahasiswa di bidang kedokteran dasar (Kedokteran) maupun kedokteran klinik, yaitu pemilihan uji statistik yang tepat untuk menganalisis data penelitian. Di era kedokteran berbasis bukti atau *evidence-based medicine*, kemampuan untuk menerjemahkan data mentah menjadi bukti ilmiah yang valid tidak lagi menjadi keahlian khusus seorang biostatistikawan, melainkan sebuah kompetensi inti bagi setiap insan medis. Kesalahan dalam memilih metode analisis tidak hanya berisiko menghasilkan kesimpulan yang keliru, tetapi juga dapat mendelegitimasi seluruh upaya penelitian, bahkan berpotensi membahayakan keselamatan pasien jika diterapkan dalam praktik klinis. Urgensi penguasaan materi ini semakin meningkat seiring dengan kompleksitas desain riset modern dan volume data yang terus bertambah.

Tujuan utama dari penulisan buku ini adalah untuk menyediakan panduan yang sistematis, logis, dan aplikatif bagi pembaca dalam menavigasi lanskap uji statistik. Buku ini dirancang untuk menjembatani kesenjangan antara pemahaman teoretis tentang statistik dan aplikasinya dalam konteks penelitian Kedokteran yang nyata. Dengan alur yang terstruktur, mulai dari konsep dasar hingga analisis multivariat lanjutan, pembaca akan dibimbing untuk membangun kerangka berpikir analitis. Kerangka ini memungkinkan mereka untuk mengidentifikasi pertanyaan penelitian, mengenali jenis data, memahami asumsi yang mendasari setiap uji, dan pada akhirnya, membuat keputusan metodologis yang dapat dipertanggungjawabkan secara ilmiah. Buku ini diharapkan dapat menjadi sumber daya esensial

yang memberdayakan peneliti untuk menghasilkan karya ilmiah yang tidak hanya signifikan secara statistik, tetapi juga relevan secara klinis.



www.penerbitbukumurah.com

**Dilarang keras, mencetak naskah
hasil layout ini tanpa seijin Penerbit**



KATA PENGANTAR

Kehadiran buku ajar *Pemilihan Uji Statistik Kedokteran* di hadapan para pembaca, khususnya mahasiswa, peneliti, dan para praktisi di bidang kesehatan, merupakan buah dari sebuah kegelisahan akademis. Seringkali, statistik dipersepsikan sebagai disiplin yang abstrak dan menakutkan, padahal ia adalah bahasa universal dari sains. Harapan terbesar penulis adalah agar buku ini dapat mengubah persepsi tersebut dan menjadikan statistik sebagai alat yang intuitif dan memberdayakan. Semoga setiap bab yang tersaji mampu menuntun pembaca untuk melihat data bukan sekadar sebagai angka, melainkan sebagai narasi yang menunggu untuk diungkapkan dengan metode yang benar. Penulis berharap buku ini tidak hanya menjadi referensi saat menyusun tugas akhir atau publikasi, tetapi juga menjadi fondasi untuk berpikir kritis dan analitis sepanjang karier profesional pembaca di dunia Kedokteran.

Penyelesaian buku ajar ini tidak akan terwujud tanpa dukungan dan kontribusi dari berbagai pihak. Ucapan terima kasih yang tulus penulis sampaikan kepada rekan-rekan sejawat yang telah memberikan masukan berharga selama proses penulisan. Apresiasi setinggi-tingginya juga ditujukan kepada para mahasiswa yang, melalui pertanyaan dan diskusi kritis di ruang kelas, secara tidak langsung telah menginspirasi penulis agar lebih relevan dan mudah dipahami. Keluarga tercinta menjadi sumber kekuatan yang tidak ternilai, memberikan dukungan moral yang tiada henti. Akhir kata, penulis menyadari bahwa buku ini masih jauh dari kesempurnaan. Oleh karena itu, segala bentuk kritik dan saran yang membangun dari para pembaca akan diterima dengan tangan terbuka demi perbaikan di masa mendatang. Selamat membaca dan semoga bermanfaat.



www.penerbitbukumurah.com

**Dilarang keras, mencetak naskah
hasil layout ini tanpa seijin Penerbit**



DAFTAR ISI

BAB

01

PRAKATA.....	i
KATA PENGANTAR.....	iii
DAFTAR ISI	v

PENGANTAR STATISTIKA DALAM RISET

KEDOKTERAN.....	1
1.1 Peran Statistik dalam Metodologi Penelitian Kedokteran.....	4
1.2 Variabel dan Skala Pengukuran Data.....	11
1.3 Alur Logika Pemilihan Uji Statistik	18
1.4 Populasi, Sampel, dan <i>Sampling</i>	24
1.5 Studi Kasus: Implementasi Logika Statistik pada Riset Klinis.....	29

DISTRIBUSI DATA DAN UJI NORMALITAS..... 41

2.1 Konsep Distribusi Normal (Gaussian)	44
2.2 Teknik Uji Normalitas Data Numerik.....	49
2.3 Uji Homogenitas Varians	54
2.4 Studi Kasus: Evaluasi Sebaran Data Parameter Laboratorium.....	59

BAB

02

STATISTIK PARAMETRIK VS NON-PARAMETRIK... 71

3.1 Prinsip Dasar Statistik Parametrik	74
3.2 Pengenalan Statistik Non-Parametrik.....	79
3.3 <i>Robustness</i> dan Kekuatan Uji	83
3.4 Studi Kasus: Dilema Pemilihan Metode pada Sampel Kecil	87

BAB

03

BAB

04

UJI KOMPARATIF DUA KELOMPOK (DATA NUMERIK) 99

- 4.1 *Independent T-Test* (Uji T Tidak Berpasangan).....102
- 4.2 *Paired T-Test* (Uji T Berpasangan).....106
- 4.3 Alternatif Non-Parametrik: Mann-Whitney dan Wilcoxon 110
- 4.4 Studi Kasus: Uji Efektivitas Antihipertensi Baru 114

BAB

05

UJI KOMPARATIF LEBIH DARI DUA KELOMPOK (ANOVA) 123

- 5.1 *One-Way ANOVA* (Analisis Varians Satu Arah).....126
- 5.2 Repeated Measures ANOVA 130
- 5.3 Alternatif Non-Parametrik: Kruskal-Wallis dan Friedman 134
- 5.4 Studi Kasus: Perbandingan Multi-Dosis Obat Analgesik.....138

BAB

06

ANALISIS DATA KATEGORIK (*CHI-SQUARE*) 147

- 6.1 Uji *Chi-Square* (X^2) Pearson 150
- 6.2 Uji Fisher *Exact* dan Yates *Correction* 153
- 6.3 Uji McNemar dan Cochran's Q.....156
- 6.4 Studi Kasus: Hubungan Merokok dengan Kejadian Kanker Paru 160

BAB

07

ANALISIS KORELASI DAN ASOSIASI 169

- 7.1 Korelasi Pearson (*Product Moment*) 172
- 7.2 Korelasi Spearman dan Kendall's Tau 175
- 7.3 Interpretasi Koefisien Determinasi (R^2).....178
- 7.4 Studi Kasus: Hubungan IMT dengan Tekanan Darah 181

BAB
08

ANALISIS REGRESI LINIER.....	191
8.1 Regresi Linier Sederhana	194
8.2 Regresi Linier Berganda (<i>Multiple Linear Regression</i>).....	197
8.3 Diagnostik Model dan <i>Residual</i>	201
8.4 Studi Kasus: Prediksi Kadar Kreatinin Berdasarkan Usia dan Berat Badan	204

BAB
09

REGRESI LOGISTIK DALAM KEDOKTERAN	213
9.1 Dasar Regresi Logistik Biner	216
9.2 Regresi Logistik Multinomial dan Ordinal.....	219
9.3 Evaluasi Model Prediktif.....	223
9.4 Studi Kasus: Prediksi Mortalitas Pasien ICU	226

BAB
10

ANALISIS <i>SURVIVAL</i> (KETAHANAN HIDUP)	235
10.1 Konsep Dasar <i>Time-to-Event</i> Data	238
10.2 Metode Kaplan-Meier dan <i>Log-Rank Test</i>	241
10.3 Model Regresi Cox <i>Proportional Hazard</i>	244
10.4 Studi Kasus: Ketahanan Hidup Pasien Pasca Transplantasi Hati	247

BAB
11

ANALISIS VALIDITAS DAN RELIABILITAS INSTRUMEN	255
11.1 Uji Validitas Konstruksi dan Isi.....	258
11.2 Uji Reliabilitas (Konsistensi Internal)	261
11.3 Uji Diagnostik: Sensitivitas dan Spesifisitas	264
11.4 Studi Kasus: Validasi Kuesioner Tingkat Kecemasan Pasien Pre-Operasi.....	267

BAB
12

ANALISIS MULTIVARIAT LANJUTAN (PENGANTAR)	275
12.1 Analisis Kovarians (ANCOVA).....	278
12.2 Analisis Multivariat Varians (MANOVA)	281
12.3 Pengantar Analisis Multivariat Lainnya	284

BAB**13**

12.4 Studi Kasus: Eksplorasi Data Kedokteran Kompleks	287
---	-----

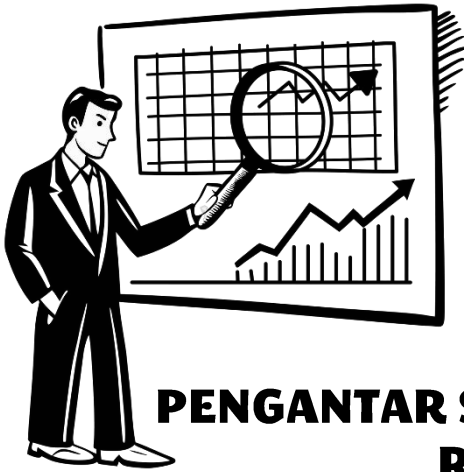
PELAPORAN HASIL STATISTIK DAN INTERPRETASI JURNAL..... 295

13.1 Standar Pelaporan Statistik (CONSORT & STROBE).....	297
13.2 Interpretasi Klinis vs Signifikansi Statistik	301
13.3 Kesalahan Umum dalam Statistik Kedokteran	304
13.4 Studi Kasus: Bedah Statistik Jurnal Kedokteran Internasional.....	307

BAB**14****WORKSHOP SOFTWARE STATISTIK (SPSS/R/ STATA)..... 315**

14.1 Manajemen Data dan <i>Cleaning</i>	318
14.2 Running Uji Statistik Populer (<i>Step-by-Step</i>) ..	321
14.3 Penutup: Memilih Uji yang Tepat untuk Tugas Akhir.....	325
14.4 Studi Kasus: <i>Workshop</i> Mandiri Pengolahan Data Riset	328

GLOSARIUM 337**REFERENSI 349****PROFIL PENULIS..... 361**



Bab 01.

PENGANTAR STATISTIKA DALAM Riset Kedokteran

Tujuan Pembelajaran

Setelah mempelajari bab ini, mahasiswa diharapkan mampu:

- Menjelaskan perbedaan fundamental antara statistik deskriptif dan inferensial serta perannya dalam metodologi riset Kedokteran.
- Mengidentifikasi pentingnya bioetika dan integritas data dalam setiap tahapan analisis statistik.
- Mengklasifikasikan variabel penelitian berdasarkan skala pengukuran data, yaitu kategorik (nominal, ordinal) dan numerik (interval, rasio).
- Menganalisis konsep akurasi, presisi, validitas, dan reliabilitas instrumen sebagai fondasi data yang berkualitas.
- Memahami alur logika sistematis dalam pemilihan uji statistik yang dimulai dari perumusan hipotesis.
- Membedakan antara populasi dan sampel, serta menjelaskan berbagai teknik sampling yang relevan untuk riset klinis.
- Menerapkan konsep dasar statistika dalam sebuah studi kasus riset klinis, mulai dari desain hingga simulasi pemilihan uji.

Pendahuluan

Pernahkah Anda mempertanyakan bagaimana sebuah obat baru dapat disetujui untuk penggunaan massal, atau bagaimana faktor risiko suatu penyakit dapat diidentifikasi dari ribuan kasus pasien? Jawaban atas pertanyaan-pertanyaan krusial dalam dunia kedokteran hampir selalu berakar pada satu disiplin ilmu yang sama, yaitu statistika. Statistik bukan sekadar kumpulan angka atau grafik, melainkan sebuah perangkat metodologis yang kuat untuk menavigasi ketidakpastian yang melekat dalam data biologis dan klinis. Dalam riset Kedokteran, statistika berfungsi sebagai tulang punggung yang menopang seluruh struktur, mulai dari perancangan studi, pengumpulan data, analisis, hingga interpretasi hasil yang menjadi landasan pengambilan keputusan klinis. Tanpa pemahaman statistik yang solid, seorang peneliti atau praktisi medis akan kesulitan untuk membedakan antara temuan yang bermakna dan kebetulan semata.

Penguasaan statistika memberikan kemampuan secara kritis untuk mengevaluasi literatur medis yang menjadi dasar dari praktik kedokteran berbasis bukti atau *Evidence-Based Medicine* (EBM). Kemampuan ini memungkinkan para profesional kesehatan untuk memahami kekuatan dan keterbatasan sebuah studi, menginterpretasikan *p-value* dan interval kepercayaan dengan benar, serta menilai apakah hasil penelitian tersebut dapat diterapkan pada populasi pasien mereka (Al-Jundi & Sakka, 2022). Oleh karena itu, literasi statistik bukan lagi sebuah pilihan, melainkan sebuah keharusan profesional. Kegagalan dalam menerapkan prinsip statistik yang benar dapat mengarah pada pengambilan kesimpulan yang salah, membuang sumber daya yang berharga, dan yang terpenting dapat berdampak negatif pada perawatan pasien.

Buku ini dirancang untuk demistifikasi statistik bagi komunitas kedokteran. Alih-alih berfokus pada formula matematis yang rumit, penekanannya adalah pada logika dan intuisi di balik pemilihan uji statistik. Pendekatan ini didasarkan pada keyakinan bahwa tujuan utama pendidikan statistik bagi peneliti kedokteran adalah untuk

menjadi pengguna statistik yang cerdas, mampu berkolaborasi secara efektif dengan ahli biostatistik dan membuat keputusan metodologis yang tepat. Memahami 'mengapa' dan 'kapan' menggunakan suatu uji seringkali lebih penting daripada mengetahui 'bagaimana' menghitungnya secara manual, terutama di era ketersediaan perangkat lunak statistik yang canggih.

Bab pertama ini akan menjadi fondasi bagi seluruh pembahasan dalam buku ini. Kita akan memulai dengan menjelajahi peran ganda statistik, baik sebagai alat untuk mendeskripsikan data (statistik deskriptif) maupun sebagai metode untuk membuat generalisasi dari sampel ke populasi (statistik inferensial). Pembahasan ini akan mencakup pentingnya integritas data dan etika penelitian sebagai prasyarat utama sebelum analisis apa pun dilakukan (Resnik, 2023). Integritas data merupakan pilar yang memastikan bahwa kesimpulan statistik yang ditarik berasal dari informasi yang akurat dan dapat dipercaya merupakan inti dari riset yang bertanggung jawab.

Selanjutnya, kita akan mendalami materi tentang 'bahan baku' dari semua analisis statistik, yaitu data itu sendiri. Pemahaman mendalam mengenai klasifikasi variabel dan skala pengukurannya adalah kunci pertama dalam alur pemilihan uji statistik (Kim et al., 2024). Kesalahan dalam mengidentifikasi tipe data akan secara otomatis mengarahkan pada pemilihan uji yang salah, tidak peduli seberapa canggih analisis yang dilakukan setelahnya. Hal ini seringkali menjadi kesalahan pemula yang paling umum namun paling fatal dalam riset Kedokteran.

Logika pemilihan uji statistik akan diperkenalkan melalui sebuah kerangka kerja yang sistematis. Kerangka ini dimulai dari langkah paling fundamental, yaitu perumusan pertanyaan penelitian dan hipotesis yang jelas (Venkatesh & Santhosh, 2025). Pertanyaan penelitian yang dirumuskan dengan baik akan secara alami menuntun peneliti pada tujuan analisis, apakah itu untuk membandingkan perbedaan, mengukur hubungan, atau memprediksi hasil. Struktur berpikir logis ini adalah kompas yang akan memandu kita melalui berbagai pilihan uji statistik yang akan dibahas di bab-bab selanjutnya.

Konsep populasi dan sampel juga akan dibahas secara mendalam, karena validitas eksternal sebuah penelitian, atau sejauh mana hasilnya dapat digeneralisasi, sangat bergantung pada bagaimana sampel dipilih dari populasi target (Berger et al., 2021). Pembahasan akan mencakup teknik-teknik *sampling* serta pentingnya kalkulasi ukuran sampel yang memadai untuk memastikan kecukupan kekuatan statistik dalam mendeteksi efek yang sebenarnya ada. Tanpa sampel yang representatif, hasil penelitian yang paling signifikan secara statistik sekalipun mungkin tidak memiliki relevansi di dunia nyata.

Sebagai penutup bab, semua konsep teoretis yang telah dipelajari akan diintegrasikan dan diterapkan dalam sebuah studi kasus komprehensif. Studi kasus ini akan mensimulasikan proses berpikir seorang peneliti klinis, mulai dari merancang protokol studi, mendefinisikan variabel, hingga menggunakan diagram alur untuk memilih uji statistik yang paling sesuai. Melalui contoh konkret ini, diharapkan pembaca dapat melihat bagaimana elemen-elemen dasar statistika ini bersatu padu untuk membentuk sebuah metodologi penelitian yang kokoh dan dapat dipertanggungjawabkan (Pugh & Torres-Saavedra, 2021). Bab ini adalah langkah awal untuk membangun intuisi statistik Anda.

1.1 Peran Statistik dalam Metodologi Penelitian Kedokteran

Dalam arsitektur penelitian kedokteran, statistika memegang peranan ganda yang krusial, yaitu sebagai fondasi sekaligus kerangka. Sebagai fondasi, statistika menyediakan prinsip dan metode untuk merancang sebuah studi yang efisien dan bebas bias, memastikan bahwa data yang dikumpulkan valid dan reliabel. Sebagai kerangka, statistika menawarkan perangkat untuk menganalisis data tersebut, mengubahnya dari sekumpulan observasi mentah menjadi bukti yang dapat diinterpretasikan dan digunakan untuk pengambilan keputusan. Tanpa perancangan statistik yang tepat, sebuah penelitian berisiko menghasilkan data yang tidak dapat dianalisis secara bermakna, sebuah kondisi yang sering diistilahkan sebagai *garbage in, garbage*

out. Metodologi penelitian yang kuat adalah hasil dari integrasi antara keahlian substansif di bidang medis dengan rigor dari prinsip-prinsip statistik (Pugh & Torres-Saavedra, 2021).

Peran statistika dimulai jauh sebelum data pertama dikumpulkan. Pada fase desain penelitian, statistika membantu peneliti dalam merumuskan hipotesis yang dapat diuji (*testable hypothesis*), menentukan desain studi yang paling sesuai (misalnya, kohort, kasus-kontrol, atau uji klinis acak terkontrol), serta menghitung ukuran sampel minimum yang diperlukan (Serdar et al., 2021). Kalkulasi ukuran sampel adalah salah satu kontribusi statistik yang paling vital, karena studi dengan sampel yang terlalu kecil (*underpowered*) mungkin gagal mendeteksi efek yang sebenarnya ada, sementara studi dengan sampel yang terlalu besar dapat menjadi tidak etis dan boros sumber daya. Setiap keputusan desain ini memiliki implikasi langsung terhadap validitas dan kekuatan kesimpulan yang akan ditarik.

Setelah data terkumpul, peran statistika beralih ke analisis dan interpretasi. Di sinilah statistik deskriptif dan inferensial bekerja secara sinergis. Statistik deskriptif digunakan untuk meringkas dan menyajikan karakteristik dasar dari data, memberikan gambaran umum mengenai sampel penelitian. Statistik inferensial kemudian mengambil langkah lebih jauh dengan menggunakan data dari sampel tersebut untuk membuat kesimpulan atau generalisasi tentang populasi yang lebih besar tempat sampel itu berasal (Kim et al., 2024). Proses inferensi ini merupakan inti dari pengujian hipotesis, di mana peneliti secara formal mengevaluasi apakah data yang diamati mendukung atau menolak hipotesis awal?

Statistika juga memainkan peran penting dalam mengelola dan mengkuantifikasi ketidakpastian. Dalam penelitian kedokteran, variabilitas biologis adalah sebuah keniscayaan. Dua pasien dengan penyakit yang sama mungkin merespons pengobatan yang sama secara berbeda. Statistik menyediakan alat, seperti interval kepercayaan (*confidence interval*) dan *p-value* untuk mengukur ketidakpastian ini (De Oliveira Santana Amaral & Line, 2021). Interval kepercayaan

memberikan rentang nilai yang masuk akal untuk parameter populasi (misalnya, rata-rata penurunan tekanan darah), sementara *p-value* membantu menilai kekuatan bukti terhadap hipotesis nol.

Integritas metodologi penelitian medis modern sangat bergantung pada penerapan statistik yang benar. Pedoman pelaporan internasional seperti CONSORT (untuk uji klinis) dan STROBE (untuk studi observasional) menempatkan penekanan besar pada pelaporan metode statistik secara transparan dan lengkap (Christensen et al., 2023). Hal ini memungkinkan pembaca dan peneliti lain untuk secara kritis menilai validitas studi dan mereplikasi temuan jika diperlukan. Transparansi ini adalah pilar dari proses peninjauan sejawat (*peer review*) dan kemajuan ilmu pengetahuan secara kumulatif.

Lebih jauh, statistika memungkinkan sintesis bukti dari berbagai studi melalui teknik meta-analisis. Dengan menggabungkan hasil dari beberapa penelitian yang relevan secara statistik, meta-analisis dapat memberikan estimasi efek yang lebih presisi dan kuat daripada studi individual mana pun (Goligher et al., 2024). Kemampuan untuk mensintesis bukti ini sangat fundamental bagi pengembangan pedoman praktik klinis dan kebijakan kesehatan, memastikan bahwa keputusan didasarkan pada totalitas bukti ilmiah yang tersedia, bukan hanya pada satu atau dua studi yang mungkin hasilnya berbeda karena kebetulan.

Akhirnya, peran statistik tidak berhenti pada publikasi hasil. Statistik juga digunakan dalam pemantauan keamanan obat (*pharmacovigilance*), evaluasi alat diagnostik, dan pemodelan prediksi penyakit. Dalam setiap aplikasi ini, statistika menyediakan bahasa dan metode yang objektif untuk mengevaluasi bukti dan membuat keputusan di tengah ketidakpastian. Oleh karena itu, pemahaman yang kuat tentang peran statistik dalam metodologi penelitian bukan hanya keterampilan teknis, tetapi juga komponen esensial dari pemikiran kritis dalam ilmu kedokteran.

Analogi: Statistik dalam riset Kedokteran dapat diibaratkan sebagai sistem peradilan dalam masyarakat. Pertanyaan penelitian adalah

terdakwa. Hipotesis nol adalah asumsi 'tidak bersalah' (tidak ada efek). Data adalah bukti-bukti yang disajikan di pengadilan. Uji statistik adalah proses persidangan itu sendiri yang secara objektif mengevaluasi kekuatan bukti. *P-value* adalah vonis hakim yang menentukan apakah bukti yang ada cukup kuat untuk menolak asumsi tidak *bersalah*. Sama seperti sistem peradilan yang bertujuan untuk mencapai keadilan berdasarkan bukti, statistika bertujuan untuk mencapai kebenaran ilmiah berdasarkan data.

1.1.1 Perbedaan Statistik Deskriptif dan Inferensial dalam Kedokteran

Statistik deskriptif adalah cabang statistika yang berfokus pada pengorganisasian, peringkasan, dan penyajian data dengan cara yang informatif. Tujuannya adalah untuk memberikan gambaran atau deskripsi ringkas tentang karakteristik utama dari kumpulan data yang ada tanpa membuat kesimpulan apa pun di luar data itu sendiri (Guzik & Więckowska, 2023). Dalam konteks Kedokteran, statistik deskriptif digunakan secara ekstensif pada bagian awal dari sebuah laporan penelitian, seringkali dalam bentuk tabel karakteristik demografis (Tabel 1) dari partisipan studi. Ukuran yang umum digunakan meliputi ukuran tendensi sentral seperti rerata (*mean*), median, dan modus untuk data numerik, serta ukuran penyebaran seperti standar deviasi, rentang interkuartil (*interquartile range*), dan rentang (*range*). Untuk data kategorik, deskripsi biasanya disajikan dalam bentuk frekuensi dan persentase.

Di sisi lain, statistik inferensial menggunakan data dari sampel penelitian untuk membuat estimasi, prediksi, atau generalisasi tentang populasi yang lebih besar (Kim et al., 2024). Ini adalah lompatan dari yang diketahui (data sampel) ke yang tidak diketahui (karakteristik populasi). Proses ini secara inheren mengandung ketidakpastian yang dikuantifikasi menggunakan teori probabilitas. Sebagai contoh, ketika sebuah uji klinis menemukan bahwa obat baru secara signifikan menurunkan tekanan darah pada 500 pasien (sampel), tujuan utamanya adalah untuk menyimpulkan (*infer*) bahwa obat tersebut juga akan

efektif untuk jutaan pasien lain dengan kondisi serupa (populasi). Alat utama statistik inferensial adalah pengujian hipotesis dan estimasi interval kepercayaan yang memungkinkan peneliti untuk menilai signifikansi statistik dari temuan mereka.

Perbedaan fundamental terletak pada cakupan kesimpulan. Kesimpulan dari statistik deskriptif hanya berlaku untuk kelompok data yang dianalisis (sampel). Jika rerata usia partisipan dalam sebuah studi adalah 55 tahun, ini adalah fakta deskriptif tentang sampel tersebut, dan hanya itu. Sebaliknya, statistik inferensial bertujuan untuk membuat klaim yang lebih luas. Dengan menggunakan data sampel yang sama, peneliti mungkin menyimpulkan dengan keyakinan 95% bahwa rerata usia populasi dari mana sampel diambil berada di antara 53 dan 57 tahun (Pugh & Torres-Saavedra, 2021). Pernyataan kedua ini adalah sebuah inferensi, sebuah generalisasi yang melampaui data yang ada dan secara eksplisit mengakui adanya *margin of error*.

1.1.2 Pentingnya Bioetika dan Integritas Data Statistik

Bioetika dan integritas data merupakan landasan moral dan ilmiah dari setiap penelitian kedokteran yang melibatkan analisis statistik. Integritas data mengacu pada keakuratan, kelengkapan, dan konsistensi data sepanjang siklus hidupnya, mulai dari pengumpulan, penyimpanan, hingga analisis (Resnik, 2023). Tanpa integritas data, analisis statistik yang paling canggih sekalipun akan menghasilkan kesimpulan yang tidak valid dan menyesatkan. Pelanggaran integritas dapat berkisar dari kesalahan yang tidak disengaja, seperti salah ketik saat entri data, hingga pelanggaran etika yang serius seperti fabrikasi (menciptakan data fiktif) atau falsifikasi (mengubah atau menghapus data yang ada agar sesuai dengan hasil diinginkan). Praktik-praktik ini tidak hanya merusak kredibilitas peneliti tetapi juga mengikis kepercayaan publik terhadap sains.

Prinsip-prinsip bioetika, seperti yang diuraikan dalam Deklarasi Helsinki memberikan kerangka kerja untuk memastikan bahwa penelitian dilakukan secara bertanggung jawab, terutama ketika melibatkan subjek manusia. Dalam konteks statistik, ini berarti bahwa desain

penelitian harus memiliki kekuatan statistik (*statistical power*) yang memadai untuk menjawab pertanyaan penelitian, karena mendaftarkan partisipan dalam studi yang *underpowered* dianggap tidak etis karena memaparkan mereka pada potensi risiko tanpa kemungkinan yang realistis untuk menghasilkan pengetahuan yang bermanfaat (Berger et al., 2021). Selain itu, semua hasil, baik yang signifikan secara statistik maupun tidak, harus dilaporkan secara transparan untuk menghindari bias publikasi, yaitu kecenderungan untuk hanya menerbitkan hasil yang positif.

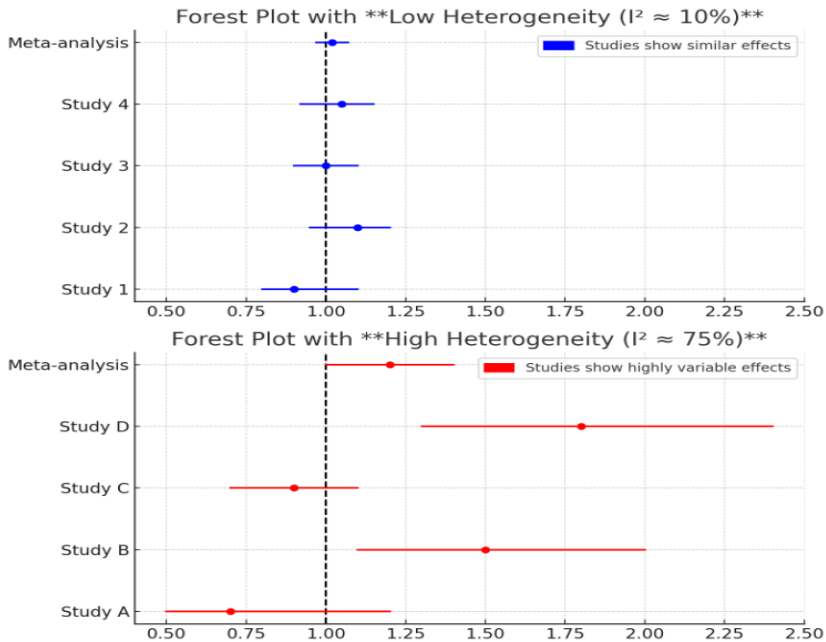
Manajemen data yang etis juga mencakup perlindungan kerahasiaan dan privasi partisipan. Data kedokteran seringkali bersifat sangat sensitif. Oleh karena itu, prosedur untuk de-identifikasi data (menghapus informasi pengenalan pribadi) dan penyimpanan data yang aman adalah wajib (Tenny & Abdelgawad, 2023). Setiap analisis statistik harus dilakukan pada data yang telah dianonimkan sejauh mungkin untuk melindungi partisipan. Persetujuan setelah penjelasan (*informed consent*) yang diperoleh dari partisipan juga harus secara jelas menyatakan bagaimana data mereka akan digunakan, dianalisis, dan dibagikan, memastikan bahwa partisipan memiliki otonomi penuh atas informasi mereka.

1.1.3 Aplikasi Statistik dalam *Evidence-Based Medicine* (EBM)

EBM adalah pendekatan dalam praktik medis yang bertujuan untuk mengintegrasikan bukti ilmiah terbaik yang tersedia dengan keahlian klinis dan nilai-nilai pasien untuk membuat keputusan perawatan kesehatan. Statistik adalah mesin yang menggerakkan EBM, menyediakan alat untuk menghasilkan, mengevaluasi, dan mensintesis bukti tersebut (Al-Jundi & Sakka, 2022). Aplikasi statistik dalam EBM sangat luas, mulai dari evaluasi uji klinis acak terkontrol (*Randomized Controlled Trials* atau RCT) yang dianggap sebagai standar emas untuk bukti efektivitas intervensi, hingga analisis studi observasional untuk mengidentifikasi faktor risiko atau prognosis. Tanpa metode statistik, mustahil untuk menentukan apakah suatu intervensi baru lebih baik daripada standar perawatan yang ada.

Salah satu aplikasi statistik yang paling kuat dalam EBM adalah meta-analisis. Ini adalah teknik statistik yang menggabungkan hasil kuantitatif dari beberapa studi independen yang membahas pertanyaan penelitian yang sama. Dengan meningkatkan ukuran sampel secara efektif, meta-analisis dapat memberikan estimasi efek pengobatan yang lebih presisi dan dapat diandalkan daripada studi individual mana pun (Goligher et al., 2024). Hasilnya seringkali disajikan secara visual dalam bentuk *forest plot* (Gambar 1.) yang memungkinkan para klinisi untuk dengan cepat memahami besaran dan konsistensi efek di berbagai studi. Meta-analisis merupakan puncak dari piramida bukti EBM dan seringkali menjadi dasar untuk pedoman praktik klinis.

Statistik juga sangat penting dalam pengembangan dan evaluasi alat diagnostik. Konsep seperti sensitivitas, spesifisitas, nilai prediksi positif (*Positive Predictive Value*), dan nilai prediksi negatif (*Negative Predictive Value*) semuanya berasal dari analisis statistik pada tabel kontingensi 2x2. Ukuran-ukuran ini membantu klinisi memahami seberapa baik sebuah tes diagnostik dapat mengidentifikasi individu dengan penyakit (sensitivitas) dan individu tanpa penyakit (spesifisitas) (Pugh & Torres-Saavedra, 2021). Kurva Karakteristik Operasi Penerima (*Receiver Operating Characteristic* atau ROC) adalah alat statistik lain yang digunakan untuk menentukan titik potong (*cut-off point*) optimal untuk uji diagnostik dan membandingkan kinerja beberapa uji.



Gambar 1. *Forest plot* dengan heterogenitas rendah (atas) dan tinggi (bawah). *Forest plot* dengan heterogenitas rendah ($I^2 \approx 10\%$) memperlihatkan ukuran efek yang konsisten (semua studi melaporkan nilai OR yang serupa sekitar 1,0). Interval kepercayaan (CI) tumpang tindih secara signifikan, menunjukkan sedikit variasi antar studi. Hasil meta-analisis stabil, artinya penggabungan studi-studi ini valid. Interpretasi bahwa heterogenitas rendah menunjukkan konsistensi yang kuat antar studi. *Forest plot* dengan heterogenitas tinggi ($I^2 \approx 75\%$) memperlihatkan ukuran efek sangat bervariasi (beberapa studi mendukung pengobatan, yang lain menunjukkan tidak ada efek atau bahkan berbahaya). Interval kepercayaan tidak tumpang tindih dengan baik, menunjukkan variabilitas besar antar studi. Ringkasan meta-analisis kurang dapat diandalkan karena hasilnya terlalu berbeda. Interpretasi bahwa heterogenitas tinggi berarti variasi yang substansial membutuhkan investigasi lebih dalam, misalnya analisis subkelompok (mygpnotes.com, 2024).

1.2 Variabel dan Skala Pengukuran Data

Memahami sifat data adalah langkah pertama yang paling fundamental sebelum melakukan analisis statistik apa pun. Setiap informasi yang dikumpulkan dalam sebuah penelitian direkam sebagai variabel, yang

secara harfiah berarti suatu atribut yang nilainya dapat bervariasi antar individu atau dari waktu ke waktu. Variabel dapat berupa karakteristik demografis (usia, jenis kelamin), pengukuran fisiologis (tekanan darah, kadar glukosa), atau hasil intervensi (sembuh, tidak sembuh). Kemampuan untuk mengklasifikasikan variabel-variabel ini dengan benar ke dalam skala pengukuran yang sesuai adalah prasyarat mutlak untuk memilih uji statistik yang tepat (Kim et al., 2024). Kesalahan dalam langkah awal ini akan merambat dan membatalkan validitas seluruh analisis berikutnya serta merupakan sebuah kesalahan fatal yang sering terjadi dalam penelitian.

Skala pengukuran yang pertama kali diusulkan oleh Stanley Smith Stevens menyediakan sebuah hierarki untuk mengklasifikasikan data. Hierarki ini terdiri atas empat tingkatan: nominal, ordinal, interval, dan rasio. Dua skala pertama (nominal dan ordinal) secara kolektif disebut sebagai data kategorik, karena mereka mengelompokkan individu ke dalam kategori-kategori. Dua skala terakhir (interval dan rasio) secara kolektif dikenal sebagai data numerik (atau kuantitatif), karena mereka mewakili kuantitas atau besaran yang terukur (Vetter, 2017). Setiap tingkatan skala memiliki sifat matematis yang unik, menentukan jenis operasi matematis dan analisis statistik apa yang dapat diterapkan padanya.

Data nominal adalah tingkat pengukuran yang paling dasar, di mana angka hanya berfungsi sebagai label untuk membedakan antarkategori tanpa urutan inheren. Contohnya termasuk golongan darah (A, B, AB, O) atau status perkawinan (menikah, lajang, cerai). Anda tidak dapat mengatakan bahwa golongan darah A lebih besar dari golongan darah B. Data ordinal adalah satu tingkat lebih tinggi, memiliki kategori yang dapat diurutkan atau diperingkat, tetapi jarak antarperingkat tidak dapat diasumsikan sama. Contoh klasik dalam kedokteran adalah stadium kanker (I, II, III, IV) atau tingkat nyeri (ringan, sedang, berat) (Guzik & Więckowska, 2023). Kita tahu bahwa stadium III lebih parah dari stadium II, tetapi kita tidak dapat mengatakan bahwa perbedaan antara stadium II dan III sama dengan perbedaan antara stadium I dan II.

Data interval memiliki sifat urutan dan jarak yang sama antar nilai, tetapi tidak memiliki titik nol absolut yang bermakna. Ini berarti bahwa nol pada skala interval tidak menunjukkan ketiadaan total dari atribut yang diukur. Contoh paling umum adalah suhu dalam Celcius atau Fahrenheit. Perbedaan antara 20°C dan 30°C sama dengan perbedaan antara 30°C dan 40°C (yaitu, 10 derajat). Namun, 0°C tidak berarti tidak ada panas sama sekali, dan 20°C tidak dua kali lebih panas dari 10°C . Operasi penjumlahan dan pengurangan valid, tetapi perkalian dan pembagian tidak valid.

Tingkat pengukuran tertinggi adalah data rasio. Skala ini memiliki semua sifat skala interval, ditambah dengan adanya titik nol absolut yang bermakna. Titik nol ini menunjukkan ketiadaan total dari kuantitas yang diukur. Sebagian besar pengukuran kedokteran berada pada skala ini, seperti tinggi badan, berat badan, tekanan darah, atau konsentrasi hemoglobin (Vetter, 2017). Akibat adanya nol absolut, semua operasi aritmatika (penjumlahan, pengurangan, perkalian, pembagian) valid. Anda dapat dengan benar mengatakan bahwa pasien dengan berat 100 kg dua kali lebih berat daripada pasien dengan berat 50 kg.

Pemahaman yang jelas tentang perbedaan ini sangat penting karena menentukan pilihan statistik deskriptif dan inferensial. Misalnya, untuk meringkas data nominal, kita hanya dapat menggunakan modus dan frekuensi. Untuk data ordinal, median dan rentang interkuartil menjadi pilihan yang lebih baik. Untuk data numerik (interval/rasio) yang terdistribusi normal, rerata dan standar deviasi adalah ukuran yang paling umum digunakan (Kim et al., 2024). Demikian pula, uji statistik yang berbeda dirancang khusus untuk jenis data yang berbeda, seperti uji *Chi-Square* untuk data kategorik dan uji-t untuk data numerik.

Terkadang dalam analisis, peneliti mungkin perlu mengubah skala variabel. Proses ini yang dikenal sebagai transformasi variabel dan harus dilakukan dengan hati-hati serta justifikasi yang kuat. Sebagai contoh, data numerik seperti usia (skala rasio) dapat diubah menjadi data kategorik (misalnya, <40 tahun, 40-60 tahun, >60 tahun) untuk tujuan analisis tertentu. Namun, proses ini selalu mengakibatkan

hilangnya informasi dan presisi dan umumnya harus dihindari kecuali jika ada alasan teoretis yang kuat.

Contoh Kasus: Dalam sebuah studi tentang faktor risiko penyakit jantung koroner, peneliti mengumpulkan data berikut dari setiap partisipan: jenis kelamin (pria/wanita), riwayat merokok (tidak pernah/mantan/aktif), tingkat kolesterol total (mg/dL), dan suhu tubuh ($^{\circ}\text{C}$). Peneliti harus mengklasifikasikan setiap variabel dengan benar. Jenis kelamin adalah data nominal. Riwayat merokok adalah data ordinal karena ada urutan (aktif > mantan > tidak pernah). Tingkat kolesterol adalah data rasio karena 0 mg/dL berarti tidak ada kolesterol sama sekali dan 200 mg/dL dua kali lebih tinggi dari 100 mg/dL. Suhu tubuh dalam Celcius adalah data interval karena 0°C bukan berarti ketiadaan panas. Klasifikasi yang benar ini akan memandu pemilihan uji statistik, misalnya, uji *Chi-Square* untuk melihat hubungan antara jenis kelamin dan riwayat merokok, dan uji korelasi Pearson antara tingkat kolesterol dan tekanan darah.

1.2.1 Klasifikasi Data: Kategorik (Nominal, Ordinal) vs Numerik (Rasio, Interval)

Klasifikasi data yang akurat adalah langkah pertama dan tidak dapat ditawar dalam alur kerja analisis statistik. Secara garis besar, data dapat dibagi menjadi dua kelompok utama, yaitu data kategorik dan data numerik (kuantitatif) (Vetter, 2017). Data kategorik merepresentasikan karakteristik atau kualitas yang mengelompokkan individu ke dalam kategori yang berbeda. Tipe data ini selanjutnya dibagi lagi menjadi nominal dan ordinal. Data nominal adalah skala yang paling sederhana, di mana nilai-nilai hanya berfungsi sebagai label tanpa urutan atau peringkat. Contoh dalam riset kedokteran meliputi kelompok perlakuan (misalnya, grup plasebo vs grup obat), status vital (hidup vs meninggal), atau genotipe (AA, Aa, aa). Operasi matematis tidak dapat dilakukan pada data ini, dan analisisnya terbatas pada penghitungan frekuensi dan proporsi (Kim et al., 2024).

Data ordinal, meskipun masih kategorik, memiliki keunggulan berupa adanya urutan atau peringkat yang melekat pada kategorinya. Hal ini

memungkinkan kategori untuk diurutkan dari yang terendah hingga yang tertinggi atau sebaliknya. Contoh umum termasuk tingkat keparahan penyakit (ringan, sedang, parah), tingkat kepuasan pasien (sangat tidak puas, tidak puas, netral, puas, sangat puas), atau status sosial ekonomi (rendah, menengah, tinggi). Meskipun kita dapat menentukan urutan, kita tidak dapat mengasumsikan bahwa interval atau jarak antar kategori adalah sama (Guzik & Więckowska, 2023). Oleh karena itu, penggunaan rerata dan standar deviasi pada data ordinal umumnya tidak tepat, dan statistik berbasis peringkat seperti median dan rentang interkuartil lebih disarankan.

Data numerik, di sisi lain merepresentasikan kuantitas yang dapat diukur dan memiliki nilai numerik yang sebenarnya. Tipe ini dibagi menjadi data interval dan rasio. Data interval memiliki urutan dan interval yang sama antarnilai, tetapi tidak memiliki titik nol absolut. Artinya, nilai nol pada skala ini tidak menunjukkan ketiadaan atribut yang diukur. Contoh klasik adalah skor *intelligence quotient* (IQ) atau suhu dalam skala Celcius. Perbedaan IQ antara 100 dan 110 sama dengan antara 110 dan 120, tetapi seseorang dengan IQ 140 tidak dua kali lebih cerdas dari seseorang dengan IQ 70 (Vetter, 2017).

Skala pengukuran tertinggi adalah data rasio, yang memiliki semua karakteristik skala interval ditambah dengan adanya titik nol absolut yang bermakna. Titik nol sejati ini memungkinkan perbandingan rasio yang valid antar pengukuran. Sebagian besar pengukuran fisik dan fisiologis dalam penelitian kedokteran, seperti tinggi badan, berat badan, konsentrasi obat dalam darah, atau jumlah sel, termasuk dalam kategori ini (Kim et al., 2024). Akibat semua operasi aritmatika valid untuk data rasio, berbagai teknik statistik, terutama metode parametrik dapat diterapkan jika asumsi lainnya terpenuhi. Memahami perbedaan halus antara keempat skala ini sangat penting karena secara langsung menentukan repertoar uji statistik yang valid untuk digunakan.

1.2.2 Transformasi Variabel dalam Penelitian Kedokteran

Transformasi variabel adalah proses penerapan fungsi matematika pada data asli untuk mengubah distribusinya. Prosedur ini seringkali

diperlukan dalam penelitian kedokteran karena banyak variabel biologis secara alami tidak mengikuti distribusi normal (berbentuk lonceng), yang merupakan asumsi dasar bagi banyak uji statistik parametrik yang kuat seperti uji-t dan ANOVA (Kumar et al., 2023). Data seperti konsentrasi sitokin, waktu kelangsungan hidup, atau jumlah salinan virus seringkali menunjukkan distribusi yang miring ke kanan (*right-skewed*). Dengan melakukan transformasi, misalnya transformasi logaritmik, akar kuadrat, atau resiprokal, data yang miring ini dapat diubah menjadi distribusi yang lebih simetris dan mendekati normal.

Tujuan utama transformasi adalah untuk memenuhi asumsi statistik sehingga memungkinkan penggunaan uji parametrik yang umumnya memiliki kekuatan statistik lebih tinggi dibandingkan uji non-parametrik sebagai alternatif. Transformasi logaritmik (misalnya, $\log_{10}(x)$ atau $\ln(x)$) adalah salah satu paling umum digunakan untuk data yang miring ke kanan. Ini efektif karena fungsi logaritma menekan nilai-nilai besar lebih kuat daripada nilai-nilai kecil sehingga mengurangi kemiringan distribusi. Selain menormalkan data, transformasi juga dapat digunakan untuk menstabilkan varians (mencapai homoskedastisitas) antarkelompok atau untuk melinearkan hubungan antardua variabel sebelum melakukan analisis regresi (Zhou et al., 2023).

Meskipun bermanfaat, keputusan untuk mentransformasi variabel harus dibuat dengan hati-hati dan tidak secara otomatis. Salah satu tantangan utama adalah interpretasi hasil. Setelah transformasi, semua analisis (misalnya, perhitungan rerata, interval kepercayaan) dilakukan pada skala yang telah diubah. Rerata dari data yang ditransformasi logaritmik, misalnya, adalah rerata geometris, bukan rerata aritmatika dari data asli. Hasil ini kemudian perlu ditransformasi balik (*back-transformed*) ke skala pengukuran asli agar dapat diinterpretasikan secara klinis, yang terkadang dapat menjadi rumit. Oleh karena itu, peneliti harus selalu melaporkan transformasi apa yang digunakan dan memberikan justifikasi yang jelas untuk keputusannya, serta memastikan interpretasi akhir tetap bermakna dalam konteks kedokteran (Christensen et al., 2023).

1.2.3 Akurasi, Presisi, Validitas, dan Reliabilitas Instrumen

Dalam penelitian kedokteran, kualitas data yang dikumpulkan sangat bergantung pada kualitas instrumen pengukuran yang digunakan. Empat konsep fundamental yang digunakan untuk mengevaluasi kualitas instrumen adalah akurasi, presisi, validitas, dan reliabilitas. Akurasi mengacu pada seberapa dekat hasil pengukuran dengan nilai yang sebenarnya atau nilai standar emas. Sebuah alat pengukur tekanan darah dianggap akurat jika rata-rata pengukurannya sangat dekat dengan tekanan darah pasien yang sebenarnya (Taherdoost, 2023). Ketidakakuratan sistematis disebut sebagai bias. Presisi, di sisi lain mengacu pada seberapa konsisten atau dapat direproduksi hasil pengukuran ketika pengukuran dilakukan berulang kali pada subjek yang sama dalam kondisi yang sama. Sebuah alat yang presisi akan memberikan hasil yang sangat mirip satu sama lain dalam setiap pengukuran, meskipun belum tentu akurat. Idealnya sebuah instrumen harus akurat dan presisi.

Validitas adalah konsep yang lebih luas yang menanyakan, Apakah instrumen ini benar-benar mengukur apa yang seharusnya diukur?. Ini adalah ukuran kebenaran atau keakuratan konseptual dari sebuah instrumen. Terdapat beberapa jenis validitas, termasuk validitas isi (*content validity*) yang menilai apakah item-item dalam kuesioner secara memadai mencakup semua aspek dari konsep yang diukur, dan validitas konstruk (*construct validity*) yang menilai sejauh mana instrumen tersebut berkorelasi dengan ukuran-ukuran lain sesuai dengan hipotesis teoretis (Heale & Twycross, 2015). Misalnya, kuesioner depresi yang valid harus menunjukkan skor yang lebih tinggi pada pasien yang didiagnosis secara klinis dengan depresi dibandingkan dengan populasi umum.

Reliabilitas berkaitan erat dengan presisi dan mengacu pada konsistensi atau stabilitas pengukuran dari waktu ke waktu, antarpemilai atau antaritem dalam sebuah skala. Reliabilitas *test-retest* mengukur stabilitas dengan memberikan tes yang sama kepada individu yang sama pada dua waktu yang berbeda. Reliabilitas *inter-rater* menilai tingkat kesepakatan antara dua atau lebih pemilai independen yang

mengamati fenomena yang sama. Konsistensi internal yang sering diukur dengan Cronbach's Alfa, menilai sejauh mana item-item dalam satu skala (misalnya, kuesioner kualitas hidup) saling terkait dan mengukur konstruk yang sama (Tavakol & Dennick, 2011). Instrumen yang tidak reliabel tidak akan pernah menjadi valid, karena jika suatu alat tidak dapat menghasilkan hasil yang konsisten ia tidak mungkin mengukur apa pun secara akurat.

1.3 Alur Logika Pemilihan Uji Statistik

Memasuki dunia analisis statistik tanpa peta yang jelas dapat menjadi pengalaman yang membingungkan, mengingat banyaknya pilihan uji yang tersedia. Namun, pemilihan uji statistik bukanlah proses acak atau berdasarkan preferensi, melainkan sebuah alur logika yang sistematis dan dapat dipertahankan. Alur ini dimulai bukan dari data, melainkan dari pertanyaan penelitian itu sendiri. Pertanyaan penelitian yang dirumuskan dengan baik adalah titik awal dan kompas yang memandu seluruh proses pengambilan keputusan metodologis (Venkatesh & Santhosh, 2025). Kesalahan paling umum yang dilakukan peneliti pemula adalah mengumpulkan data terlebih dahulu dan baru kemudian memikirkan bagaimana cara menganalisisnya, sebuah pendekatan yang seringkali berakhir pada analisis yang tidak tepat atau bahkan tidak mungkin dilakukan.

Langkah pertama dalam alur logika ini adalah menerjemahkan pertanyaan penelitian menjadi hipotesis statistik yang spesifik dan dapat diuji. Ini melibatkan perumusan hipotesis nol (H_0), yang biasanya menyatakan tidak ada perbedaan atau tidak ada hubungan dan hipotesis alternatif (H_1) yang menyatakan adanya perbedaan atau hubungan (Thabane et al., 2021). Kejelasan dalam mendefinisikan hipotesis ini sangat penting karena akan secara langsung menentukan tujuan dari uji statistik yang akan dilakukan. Apakah tujuan utamanya adalah untuk membandingkan rata-rata antar kelompok, menguji hubungan antar variabel, atau memprediksi suatu hasil?

Setelah tujuan uji diidentifikasi, langkah kedua adalah memeriksa secara teliti desain penelitian dan sifat data yang telah dikumpulkan. Ini melibatkan beberapa pertanyaan kunci. Pertama, berapa banyak kelompok sampel yang akan dibandingkan? Apakah ada dua kelompok (misalnya, perlakuan vs kontrol) atau lebih dari dua kelompok? Kedua, apakah sampel-sampel tersebut independen (berasal dari individu yang berbeda) atau berpasangan (berasal dari individu yang sama pada waktu yang berbeda, misalnya, sebelum dan sesudah intervensi)? (Chicco et al., 2025). Jawaban atas pertanyaan-pertanyaan ini akan secara signifikan mempersempit pilihan uji statistik yang mungkin dilakukan.

Langkah ketiga dan yang tidak kalah pentingnya adalah mengidentifikasi skala pengukuran dari variabel yang terlibat, baik variabel independen maupun dependen. Seperti yang telah dibahas sebelumnya, apakah data tersebut kategorik (nominal, ordinal) atau numerik (interval, rasio)? (Kim et al., 2024). Aturan dasarnya adalah, jenis data dari variabel dependen (variabel hasil) seringkali menjadi penentu utama dalam memilih keluarga uji statistik yang tepat. Misalnya, jika variabel hasil adalah numerik, maka keluarga uji-t atau ANOVA mungkin relevan. Jika variabel hasil adalah kategorik, maka keluarga uji *Chi-Square* atau regresi logistik menjadi kandidat utama.

Langkah keempat dalam alur logika ini adalah memeriksa apakah data memenuhi asumsi dari uji statistik yang dipertimbangkan, terutama untuk uji parametrik. Asumsi yang paling umum adalah normalitas distribusi data. Jika data numerik terdistribusi secara normal, uji parametrik (seperti uji-t) dapat digunakan karena memiliki kekuatan statistik yang lebih tinggi. Namun, jika asumsi normalitas dilanggar, atau jika ukuran sampel sangat kecil, maka alternatif non-parametrik (seperti uji Mann-Whitney U) menjadi pilihan yang lebih tepat dan aman (Guzik & Więckowska, 2023). Pemeriksaan asumsi ini adalah langkah validasi yang krusial untuk memastikan hasil uji statistik dapat dipercaya.

Banyak buku teks dan sumber daya statistik menyajikan diagram alur atau *flowchart* untuk membantu memvisualisasikan proses pengambilan keputusan ini. Diagram alur ini memandu pengguna melalui serangkaian pertanyaan ya/tidak tentang tujuan penelitian, desain studi, dan jenis data yang pada akhirnya mengarah pada rekomendasi uji statistik yang paling sesuai (Venkatesh & Santhosh, 2025). Meskipun diagram alur ini sangat membantu, penting untuk diingat bahwa itu adalah panduan, bukan aturan yang kaku. Konteks penelitian dan pengetahuan substansif di bidangnya tetap memainkan peran penting.

Dalam praktiknya, alur logika ini seringkali bersifat iteratif. Peneliti mungkin perlu kembali ke langkah-langkah sebelumnya jika mereka menemukan masalah pada tahap analisis data, seperti pelanggaran asumsi yang parah. Pemahaman yang kokoh tentang alur logika ini tidak hanya memberdayakan peneliti untuk memilih uji yang benar tetapi juga untuk memberikan justifikasi yang kuat atas pilihan metodologis mereka dalam publikasi ilmiah dan presentasi.

Analogi: Memilih uji statistik dapat diibaratkan seperti seorang dokter yang memilih resep obat untuk pasien. Dokter tidak memberikan obat secara acak. Pertama, dokter melakukan anamnesis untuk memahami pertanyaan penelitian (keluhan utama pasien) dan merumuskan hipotesis (diagnosis banding). Kedua, dokter memeriksa desain studi (apakah ini kasus akut atau kronis, apakah ada riwayat pengobatan sebelumnya?). Ketiga, dokter melihat jenis data (hasil pemeriksaan fisik dan laboratorium, seperti suhu tubuh yang numerik atau tingkat kesadaran yang ordinal). Keempat, dokter mempertimbangkan asumsi (apakah pasien memiliki alergi terhadap obat tertentu?). Hanya setelah melalui alur logika ini, dokter dapat dengan percaya diri memilih resep (uji statistik) yang paling tepat dan efektif untuk situasi spesifik tersebut.

1.3.1 Identifikasi Pertanyaan Penelitian dan Hipotesis

Fondasi dari setiap analisis statistik yang bermakna adalah pertanyaan penelitian yang jelas, terfokus, dan dapat dijawab. Pertanyaan

penelitian berfungsi sebagai pemandu utama yang menentukan semua aspek metodologis berikutnya, termasuk desain studi, variabel yang akan diukur, dan akhirnya, uji statistik yang akan digunakan (Thabane et al., 2021). Kerangka kerja yang umum digunakan untuk merumuskan pertanyaan penelitian klinis yang baik adalah PICO, yang merupakan singkatan dari *Population/Patient*, *Intervention*, *Comparison*, dan *Outcome*. Sebagai contoh, pertanyaan yang diformulasikan dengan baik adalah: Pada pasien dewasa (P) dengan hipertensi derajat I, apakah pemberian obat A (I) lebih efektif dalam menurunkan tekanan darah sistolik (O) dibandingkan dengan plasebo (C) setelah 12 minggu pengobatan?

Setelah pertanyaan penelitian ditetapkan, langkah selanjutnya adalah menerjemahkannya ke dalam hipotesis statistik. Hipotesis adalah pernyataan formal yang dapat diuji tentang hubungan antarvariabel atau perbedaan antarkelompok. Selalu ada dua hipotesis yang saling melengkapi: hipotesis nol (H_0) dan hipotesis alternatif (H_1) (Venkatesh & Santhosh, 2025). Hipotesis nol adalah pernyataan *status quo* yang biasanya mengasumsikan tidak ada efek, tidak ada perbedaan, atau tidak ada hubungan. Untuk contoh di atas, H_0 akan menyatakan: Tidak ada perbedaan dalam rata-rata penurunan tekanan darah sistolik antara kelompok yang menerima obat A dan kelompok yang menerima plasebo.

Hipotesis alternatif (H_1) adalah kebalikan dari H_0 dan merupakan pernyataan yang ingin dibuktikan oleh peneliti. H_1 dapat bersifat dua arah (*two-tailed*) yang hanya menyatakan adanya perbedaan (misalnya, Ada perbedaan dalam rata-rata penurunan tekanan darah), atau satu arah (*one-tailed*) yang menentukan arah perbedaan (misalnya, Rata-rata penurunan tekanan darah pada kelompok obat A lebih besar daripada kelompok plasebo). Seluruh proses pengujian statistik pada dasarnya adalah prosedur untuk mengumpulkan bukti dari data sampel untuk menentukan apakah bukti tersebut cukup kuat untuk menolak hipotesis nol demi hipotesis alternatif (Pugh & Torres-Saavedra, 2021).

1.3.2 Menentukan Tujuan Uji: Komparatif, Korelatif, atau Prediktif

Berdasarkan pertanyaan penelitian dan hipotesis, tujuan utama dari analisis statistik dapat diklasifikasikan ke dalam tiga kategori besar: komparatif, korelatif, atau prediktif. Memahami kategori mana yang paling sesuai dengan pertanyaan penelitian adalah langkah penting dalam mempersempit pilihan uji statistik (Venkatesh & Santhosh, 2025). Tujuan komparatif berfokus pada perbandingan suatu hasil antara dua atau lebih kelompok. Pertanyaan penelitian yang diawali dengan Apakah ada perbedaan? atau Apakah kelompok A lebih baik dari kelompok B? biasanya mengarah pada uji komparatif. Contohnya termasuk membandingkan efektivitas dua obat, atau membandingkan kadar biomarker antara kelompok sehat dan kelompok sakit. Uji-t, ANOVA, dan *Chi-Square* adalah contoh klasik dari uji komparatif.

Tujuan korelatif atau asosiatif bertujuan untuk mengeksplorasi dan mengukur kekuatan serta arah hubungan antara dua atau lebih variabel dalam satu kelompok subjek. Pertanyaan penelitian yang umum adalah Apakah ada hubungan antara variabel X dan variabel Y? Sebagai contoh, seorang peneliti mungkin ingin mengetahui apakah ada hubungan antara IMT dan tekanan darah, atau antara tingkat pendidikan dan kepatuhan minum obat (Kim et al., 2024). Uji yang digunakan dalam analisis korelasional, seperti korelasi Pearson atau Spearman menghasilkan koefisien korelasi yang menunjukkan seberapa erat kedua variabel tersebut bergerak bersama, tetapi penting untuk diingat bahwa korelasi tidak menyiratkan sebab-akibat.

Tujuan prediktif mengambil langkah lebih jauh dari sekadar mengidentifikasi hubungan dengan tujuan untuk membangun model statistik yang dapat memprediksi nilai dari satu variabel (variabel dependen atau hasil) berdasarkan nilai dari satu atau lebih variabel lain (variabel independen atau prediktor). Pertanyaan kunci di sini adalah Faktor-faktor apa yang dapat memprediksi terjadinya suatu penyakit? atau Seberapa akurat kita dapat memprediksi skor Y dari skor X? Analisis regresi, baik regresi linier untuk hasil numerik maupun regresi logistik untuk hasil kategorik adalah alat utama untuk tujuan prediktif (Kumar et al., 2023). Model-model ini tidak hanya mengidentifikasi

prediktor yang signifikan tetapi juga mengkuantifikasi besarnya pengaruh setiap prediktor terhadap hasil.

1.3.3 Mengenal Diagram Alur (Flowchart) Pemilihan Uji

Diagram alur atau *flowchart* adalah alat bantu visual yang sangat berguna untuk memandu peneliti, terutama pemula, melalui proses logis pemilihan uji statistik. Diagram ini menyajikan serangkaian pertanyaan keputusan yang terstruktur, biasanya dalam format biner (ya/tidak), secara sistematis mempersempit pilihan uji yang mungkin berdasarkan karakteristik penelitian (Venkatesh & Santhosh, 2025). Penggunaan diagram alur dapat mengurangi kemungkinan pemilihan uji yang tidak tepat dan berfungsi sebagai daftar periksa untuk memastikan semua faktor relevan telah dipertimbangkan. Ini adalah representasi grafis dari alur logika yang dibahas sebelumnya.

Diagram alur yang tipikal biasanya dimulai dengan pertanyaan paling umum tentang tujuan utama penelitian (misalnya, komparasi vs korelasi). Dari sana, cabang-cabang diagram akan menanyakan tentang jumlah kelompok yang dibandingkan (satu, dua, atau lebih dari dua), apakah kelompok tersebut independen atau berpasangan, dan apa skala pengukuran dari variabel dependen dan independen (Kim et al., 2024). Dengan mengikuti jalur yang sesuai dengan desain penelitian, pengguna akan diarahkan ke kotak akhir yang merekomendasikan uji statistik yang paling sesuai, seperti *Independent T-Test* atau *Kruskal-Wallis Test*.

Meskipun sangat membantu, penting untuk menyadari keterbatasan diagram alur. Diagram ini seringkali merupakan penyederhanaan dan mungkin tidak mencakup semua skenario penelitian yang kompleks atau uji statistik yang lebih canggih. Selain itu, diagram alur biasanya mengarahkan ke pilihan uji, tetapi tidak secara otomatis memeriksa apakah asumsi dari uji tersebut (misalnya, normalitas, homogenitas varians) terpenuhi. Oleh karena itu, diagram alur harus digunakan sebagai titik awal atau alat bantu pengajaran, bukan sebagai pengganti pemahaman konseptual yang solid tentang prinsip-prinsip di balik setiap uji statistik dan asumsi yang mendasarinya (Christensen et al.,

2023). Penilaian akhir dan justifikasi pemilihan uji tetap menjadi tanggung jawab peneliti.

1.4 Populasi, Sampel, dan *Sampling*

Dalam setiap penelitian, tujuan utamanya adalah untuk menarik kesimpulan tentang kelompok besar individu yang dikenal sebagai populasi. Populasi adalah keseluruhan kumpulan individu, objek, atau pengukuran memiliki karakteristik umum yang diminati oleh peneliti. Dalam konteks kedokteran, populasi bisa berupa semua pasien diabetes tipe 2 di Indonesia atau semua sel kanker payudara jenis HER2-positif. Namun, dalam praktiknya hampir tidak mungkin untuk mempelajari setiap anggota populasi karena keterbatasan waktu, biaya, dan sumber daya. Oleh karena itu, peneliti mempelajari sebagian kecil dari populasi yang disebut dengan sampel (Berger et al., 2021). Sampel adalah subkelompok yang dipilih dari populasi yang lebih besar.

Proses pemilihan sampel dari populasi ini disebut sebagai *sampling*. Kualitas dari proses *sampling* ini sangat menentukan validitas eksternal atau generalisasi dari temuan penelitian. Artinya, sejauh mana kesimpulan yang ditarik dari sampel dapat secara akurat diterapkan kembali ke populasi asal. Jika sampel tidak dipilih dengan cara yang benar, ia mungkin tidak mewakili populasi secara akurat, suatu kondisi yang dikenal sebagai bias seleksi. Sampel yang bias dapat menghasilkan estimasi secara sistematis terlalu tinggi atau terlalu rendah dari parameter populasi sebenarnya sehingga membuat hasil penelitian menjadi tidak dapat diandalkan (Pugh & Torres-Saavedra, 2021).

Terdapat dua kategori utama teknik *sampling*, yaitu *sampling* probabilitas dan *sampling* non-probabilitas. Dalam *sampling* probabilitas, setiap anggota populasi memiliki peluang yang diketahui dan tidak nol untuk dipilih menjadi bagian dari sampel. Metode ini dianggap sebagai standar emas karena kemampuannya untuk menghasilkan sampel representatif dan memungkinkan penggunaan statistik inferensial untuk membuat generalisasi yang valid. Teknik-

teknik di bawah kategori ini termasuk *simple random sampling*, *stratified random sampling*, *systematic sampling*, dan *cluster sampling* (Serdar et al., 2021). Pemilihan metode probabilitas yang spesifik bergantung pada struktur populasi dan tujuan penelitian.

Di sisi lain, *sampling* non-probabilitas tidak melibatkan pemilihan acak. Sebaliknya, sampel dipilih berdasarkan kemudahan akses (*sampling* kemudahan atau *convenience sampling*) atau berdasarkan penilaian subjektif dari peneliti (*sampling* bertujuan atau *purposive sampling*). Meskipun metode ini lebih mudah dan lebih murah untuk diimplementasikan, mereka memiliki risiko bias yang jauh lebih tinggi dan secara teknis tidak memungkinkan untuk generalisasi statistik ke populasi yang lebih luas (Tenny & Abdelgawad, 2023). Meskipun demikian, teknik ini sering digunakan dalam penelitian eksplorasi, kualitatif atau ketika kerangka *sampling* populasi tidak tersedia.

Selain cara pemilihan sampel, penentuan ukuran sampel yang memadai adalah langkah krusial lainnya dalam proses *sampling*. Ukuran sampel yang terlalu kecil dapat menyebabkan studi kekurangan kekuatan statistik (*under powered*), artinya ia tidak memiliki kemampuan yang cukup untuk mendeteksi perbedaan atau hubungan yang benar-benar ada di populasi (kesalahan Tipe II). Sebaliknya, ukuran sampel yang lebih besar dari yang diperlukan dapat membuang-buang sumber daya dan dalam kasus uji klinis secara tidak etis mengekspos lebih banyak partisipan daripada yang dibutuhkan untuk menjawab pertanyaan penelitian (Serdar et al., 2021). Perhitungan ukuran sampel adalah proses statistik formal yang mempertimbangkan tingkat signifikansi yang diinginkan (α), kekuatan statistik (β), variabilitas data, dan ukuran efek yang diharapkan.

Untuk mendefinisikan populasi target secara lebih spesifik dan memastikan sampel lebih homogen, peneliti menetapkan kriteria inklusi dan eksklusi. Kriteria inklusi adalah serangkaian karakteristik yang harus dimiliki oleh calon subjek untuk dapat dimasukkan dalam penelitian. Sebaliknya, kriteria eksklusi adalah serangkaian karakteristik yang menyebabkan calon subjek dikeluarkan dari penelitian (Berger et

al., 2021). Penetapan kriteria ini dengan jelas sangat penting untuk memastikan keamanan partisipan, meningkatkan validitas internal penelitian, dan memastikan bahwa hasil penelitian dapat direplikasi.

Pada akhirnya, hubungan antara populasi, sampel, dan *sampling* adalah inti dari statistik inferensial. Kita menggunakan data dari sampel untuk menghitung statistik (misalnya, rerata sampel). Kemudian, kita menggunakan statistik sampel ini untuk membuat inferensi atau estimasi tentang parameter populasi (misalnya, rerata populasi) yang tidak diketahui. Tingkat kepercayaan kita pada inferensi ini sangat bergantung pada seberapa baik sampel tersebut mewakili populasi yang pada gilirannya ditentukan oleh metode *sampling* yang digunakan dan memadainya ukuran sampel.

Contoh Kasus: Seorang peneliti ingin mengetahui prevalensi anemia pada ibu hamil di sebuah provinsi. Populasi target adalah semua ibu hamil yang tinggal di provinsi tersebut. Mengumpulkan data dari semua ibu hamil di provinsi itu tidak memungkinkan. Oleh karena itu, peneliti memutuskan untuk mengambil sampel. Menggunakan teknik *stratified random sampling*, peneliti pertama-tama membagi provinsi menjadi strata berdasarkan kabupaten/kota. Kemudian, dari setiap kabupaten/kota, sejumlah klinik kesehatan ibu dan anak dipilih secara acak. Akhirnya, dari setiap klinik yang terpilih, sejumlah ibu hamil dipilih secara acak untuk diukur kadar hemoglobinnya. Dengan menggunakan metode *sampling* probabilitas ini, peneliti dapat memperoleh sampel sebanyak 800 orang yang representatif terhadap populasi ibu hamil di provinsi tersebut, dan hasilnya dapat digeneralisasi dengan tingkat kepercayaan tertentu.

1.4.1 Teknik Sampling pada Riset Eksperimental dan Observasional

Pemilihan teknik *sampling* sangat dipengaruhi oleh desain penelitian, yang secara umum dapat dikategorikan sebagai eksperimental atau observasional. Dalam riset eksperimental, seperti uji klinis acak terkontrol (RCT), peneliti secara aktif memberikan intervensi dan mengontrol alokasi subjek ke dalam kelompok perlakuan. Tujuan

utamanya adalah untuk membangun hubungan sebab-akibat. Dalam konteks ini, teknik *sampling* probabilitas seperti *simple random sampling* adalah ideal untuk memilih partisipan dari populasi target untuk memastikan generalisasi. Namun, yang lebih krusial dalam RCT adalah proses alokasi acak (*random assignment*) partisipan yang telah terpilih ke dalam kelompok intervensi atau kontrol, karena inilah yang meminimalkan bias dan memungkinkan atribusi kausalitas (Berger et al., 2021).

Dalam riset observasional (misalnya, studi kohort, kasus-kontrol, atau potong lintang), peneliti tidak melakukan intervensi tetapi hanya mengamati paparan dan hasil yang terjadi secara alami. Akibatnya seringkali tidak praktis atau tidak etis untuk melakukan eksperimen, studi ini penting untuk mengidentifikasi faktor risiko dan asosiasi. Teknik *sampling* probabilitas seperti *stratified* atau *cluster sampling* sering digunakan dalam studi potong lintang atau kohort berbasis populasi untuk memastikan sampel yang representatif (Pugh & Torres-Saavedra, 2021). Sebaliknya, dalam studi kasus-kontrol, peneliti seringkali menggunakan *sampling* non-probabilitas. Kasus (individu dengan penyakit) seringkali dipilih dari register rumah sakit, sementara kontrol (individu tanpa penyakit) dipilih agar sesuai (*matched*) dengan kasus berdasarkan variabel tertentu seperti usia dan jenis kelamin.

1.4.2 Menentukan Ukuran Sampel (*Sample Size*) yang Representatif

Penentuan ukuran sampel yang memadai adalah salah satu aspek paling penting dalam perencanaan penelitian kedokteran. Ukuran sampel yang representatif bukan hanya tentang jumlah, tetapi tentang memiliki jumlah subjek yang cukup untuk mencapai kekuatan statistik (*statistical power*) yang diinginkan. Kekuatan statistik adalah probabilitas untuk mendeteksi efek (perbedaan atau hubungan) yang benar-benar ada di populasi, dan secara konvensional ditetapkan pada 80% atau lebih tinggi (Serdar et al., 2021). Studi yang *underpowered* (kekuatan rendah) karena ukuran sampel yang terlalu kecil berisiko

menghasilkan temuan negatif palsu (kesalahan Tipe II), di mana peneliti gagal menolak hipotesis nol padahal sebenarnya salah.

Perhitungan ukuran sampel formal memerlukan spesifikasi beberapa parameter kunci. Pertama adalah tingkat signifikansi (α), yaitu probabilitas melakukan kesalahan Tipe I (menolak H_0 padahal benar), yang biasanya ditetapkan pada 0,05. Kedua adalah kekuatan yang diinginkan ($1 - \beta$), biasanya 0,80. Ketiga adalah estimasi variabilitas data (misalnya, standar deviasi), yang sering diperoleh dari penelitian sebelumnya atau studi percontohan (*pilot study*). Keempat, dan yang paling menantang adalah penentuan ukuran efek (*effect size*) terkecil yang dianggap penting secara klinis (Kim et al., 2024). Ukuran efek adalah besaran perbedaan atau hubungan yang ingin dideteksi oleh peneliti. Ukuran efek yang lebih kecil atau variabilitas yang lebih besar akan memerlukan ukuran sampel yang lebih besar untuk mencapai kekuatan yang sama.

1.4.3 Kriteria Inklusi dan Eksklusi dalam Riset Klinis

Kriteria inklusi dan eksklusi adalah seperangkat aturan yang mendefinisikan siapa yang memenuhi syarat untuk berpartisipasi dalam sebuah studi penelitian. Kriteria inklusi adalah karakteristik yang harus dimiliki oleh calon partisipan, sementara kriteria eksklusi adalah karakteristik yang mendiskualifikasi mereka (Berger et al., 2021). Tujuan utama dari penetapan kriteria ini adalah untuk mendefinisikan populasi studi secara spesifik, menciptakan kelompok sampel yang relatif homogen, dan memastikan keamanan partisipan. Kehomogenan sampel membantu mengurangi variabilitas yang tidak diinginkan dalam data, sehingga meningkatkan kemampuan studi untuk mendeteksi efek dari intervensi (validitas internal).

Kriteria inklusi biasanya mencakup faktor demografis (misalnya, usia, jenis kelamin), karakteristik klinis (misalnya, diagnosis penyakit tertentu, tingkat keparahan), dan lokasi geografis. Contoh kriteria inklusi adalah pria dan wanita berusia 18-65 tahun yang didiagnosis dengan diabetes melitus tipe 2 setidaknya selama satu tahun. Kriteria eksklusi dirancang untuk mengeluarkan individu yang partisipasinya dapat

membahayakan keselamatan mereka, yang mungkin tidak dapat menyelesaikan protokol studi, atau yang memiliki kondisi penyerta (*comorbidities*) yang dapat mengacaukan (*confound*) hasil penelitian (Pugh & Torres-Saavedra, 2021). Contoh kriteria eksklusi adalah pasien dengan riwayat penyakit ginjal kronis, wanita hamil atau menyusui, atau partisipasi dalam studi klinis lain dalam 30 hari terakhir. Kriteria ini harus didefinisikan secara jelas dan diterapkan secara konsisten untuk memastikan integritas dan replikabilitas penelitian.

1.5 Studi Kasus: Implementasi Logika Statistik pada Riset Klinis

Untuk mengilustrasikan bagaimana konsep-konsep yang telah dibahas dalam bab ini diterapkan dalam praktik, mari kita pertimbangkan sebuah skenario riset klinis. Sebuah perusahaan farmasi telah mengembangkan obat baru, *CardioNorm*, yang dirancang untuk menurunkan tekanan darah sistolik (TDS) pada pasien dengan hipertensi esensial. Mereka ingin melakukan uji coba klinis Fase III untuk membandingkan efektivitas dan keamanan *CardioNorm* dengan plasebo. Skenario ini akan menjadi kerangka kerja untuk menerapkan alur logika statistik, mulai dari desain hingga pemilihan uji. Ini adalah contoh klasik di mana penerapan prinsip statistik yang ketat sangat penting untuk menghasilkan bukti yang kredibel bagi regulator dan komunitas medis (Pugh & Torres-Saavedra, 2021).

Pertanyaan penelitian utama yang dirumuskan menggunakan kerangka PICO adalah: Pada pasien dewasa dengan hipertensi esensial (P), apakah pengobatan dengan *CardioNorm* (I) menghasilkan penurunan TDS yang lebih besar (O) dibandingkan dengan plasebo (C) setelah 12 minggu? Pertanyaan ini secara alami mengarah pada hipotesis. Hipotesis nol (H_0) adalah tidak ada perbedaan dalam rata-rata penurunan TDS antara kelompok *CardioNorm* dan kelompok plasebo. Hipotesis alternatif (H_1) adalah ada perbedaan dalam rata-rata penurunan TDS antara kedua kelompok tersebut (Thabane et al., 2021).

Berdasarkan pertanyaan penelitian ini, tujuan utama analisis adalah komparatif, yaitu membandingkan hasil antara dua kelompok. Desain penelitian yang dipilih adalah uji klinis acak terkontrol (*randomized controlled trial*, RCT) dengan desain kelompok paralel, di mana satu kelompok menerima *CardioNorm* dan kelompok lainnya menerima plasebo. Terdapat dua kelompok independen. Variabel hasil utama (variabel dependen) adalah perubahan TDS dari awal (*baseline*) hingga minggu ke-12. Perubahan TDS dihitung sebagai (TDS awal - TDS minggu ke-12) dan merupakan data numerik dengan skala rasio, karena memiliki nol absolut yang bermakna atau tidak ada perubahan (Kim et al., 2024).

Dengan informasi ini, kita dapat mulai menavigasi alur pemilihan uji statistik. Tujuan ujinya adalah komparatif. Ada dua kelompok yang dibandingkan (*CardioNorm* vs plasebo). Kelompok-kelompok ini independen karena terdiri atas pasien yang berbeda. Variabel dependen (perubahan TDS) adalah numerik. Berdasarkan parameter ini, kandidat utama untuk uji statistik adalah Uji-t untuk sampel independen (*Independent T-Test*) (Chicco et al., 2025). Uji ini secara spesifik dirancang untuk membandingkan rerata dari dua kelompok independen pada variabel hasil numerik.

Namun, sebelum finalisasi pemilihan uji, peneliti harus mempertimbangkan asumsi dari *Independent T-Test*. Asumsi utamanya adalah bahwa data perubahan TDS pada kedua kelompok harus terdistribusi secara normal dan varians dari kedua kelompok harus homogen (sama) (Guzik & Więckowska, 2023). Oleh karena itu, setelah data dikumpulkan, peneliti harus melakukan uji normalitas (misalnya, Shapiro-Wilk) dan uji homogenitas varians (misalnya, Uji Levene) terlebih dahulu. Jika kedua asumsi ini terpenuhi, maka *Independent T-Test* adalah pilihan yang tepat.

Jika asumsi normalitas dilanggar, alur logika akan mengarahkan peneliti ke alternatif non-parametrik, yaitu Uji Mann-Whitney U. Uji ini juga membandingkan dua kelompok independen tetapi bekerja pada peringkat (*rank*) data, bukan nilai sebenarnya sehingga tidak

memerlukan asumsi distribusi normal (Kim et al., 2024). Jika asumsi normalitas terpenuhi tetapi asumsi homogenitas varians dilanggar, peneliti dapat menggunakan variasi dari uji-t yang disebut *Welch's t-test*, yang tidak mengasumsikan varians yang sama.

Selain variabel hasil utama, penelitian ini juga akan mengumpulkan data lain, seperti data demografis (usia, jenis kelamin) dan data keamanan (kejadian efek samping, diklasifikasikan sebagai ya/tidak). Untuk membandingkan proporsi pasien yang mengalami efek samping antara kelompok *CardioNorm* dan plasebo (dua variabel kategorik, yaitu kelompok perlakuan dan kejadian efek samping), uji statistik yang sesuai adalah Uji *Chi-Square* atau, jika ukuran sampel kecil, Uji Eksak Fisher. Studi kasus ini menunjukkan bahwa satu penelitian seringkali memerlukan beberapa jenis uji statistik yang berbeda, tergantung pada pertanyaan spesifik dan jenis data yang dianalisis.

Contoh Kasus: Dalam protokol studi *CardioNorm*, peneliti menetapkan kriteria inklusi (usia 18-75 tahun, diagnosis hipertensi esensial, TDS awal antara 140-159 mmHg) dan kriteria eksklusi (hipertensi sekunder, riwayat infark miokard dalam 6 bulan terakhir, gangguan ginjal berat). Berdasarkan perhitungan ukuran sampel dengan asumsi kekuatan 90% untuk mendeteksi perbedaan klinis yang bermakna sebesar 5 mmHg dalam penurunan TDS, dibutuhkan 250 pasien per kelompok sehingga total pasien direkrut adalah 500 orang. Pasien-pasien ini kemudian dialokasikan secara acak untuk menerima *CardioNorm* atau plasebo. Di akhir studi, analisis akan dimulai dengan statistik deskriptif untuk meringkas karakteristik dasar kedua kelompok, diikuti oleh analisis inferensial menggunakan *Independent T-Test* (atau alternatifnya) untuk menguji hipotesis utama.

1.5.1 Analisis Desain Penelitian pada Uji Coba Klinis Fase III

Uji coba klinis Fase III merupakan tahap penting dalam pengembangan obat, yang dirancang untuk memberikan bukti definitif tentang efektivitas dan keamanan suatu intervensi baru pada populasi pasien yang besar dan beragam. Analisis desain pada studi *CardioNorm* menunjukkan penggunaan standar emas dalam penelitian klinis: desain

uji klinis acak, terkontrol plasebo, dan tersamar ganda (*randomized, double-blind, placebo-controlled trial*) (Berger et al., 2021). Acak berarti setiap partisipan memiliki kesempatan yang sama untuk dialokasikan ke kelompok *CardioNorm* atau plasebo yang meminimalkan bias seleksi dan memastikan bahwa kelompok-kelompok tersebut sebanding pada awal studi. Tersamar ganda, berarti baik partisipan maupun peneliti/dokter tidak mengetahui perlakuan mana yang diterima oleh setiap partisipan dapat mengurangi bias kinerja dan bias observasi.

Penggunaan kelompok kontrol plasebo sangat penting untuk membedakan efek farmakologis sebenarnya dari obat dari efek plasebo dan perubahan alami penyakit. Desain kelompok paralel di mana setiap kelompok menerima satu perlakuan secara bersamaan selama periode studi adalah desain yang paling umum dan lugas untuk studi Fase III. Durasi 12 minggu dipilih berdasarkan pengetahuan farmakokinetik obat dan waktu yang diharapkan untuk mencapai efek antihipertensi yang stabil (Pugh & Torres-Saavedra, 2021). Desain yang kuat ini secara langsung mendukung tujuan analisis komparatif dan memungkinkan inferensi kausal yang kuat jika ditemukan perbedaan yang signifikan secara statistik.

www.penerbitbukumurah.com
Dilarang keras, mencetak naskah

1.5.2 Penentuan Skala Data pada Kuesioner Kualitas Hidup Pasien

Selain mengukur tekanan darah, studi *CardioNorm* juga dapat menyertakan hasil sekunder seperti kualitas hidup (*Quality of Life/QoL*) yang sering diukur menggunakan kuesioner terstandardisasi, seperti SF-36 atau EQ-5D. Penentuan skala data dari kuesioner ini memerlukan analisis yang cermat. Sebagian besar item dalam kuesioner QoL menggunakan skala Likert, misalnya, meminta pasien untuk menilai kesehatan mereka pada skala dari 1 (sangat buruk) hingga 5 (sangat baik). Secara teknis, data ini adalah data ordinal karena meskipun ada urutan yang jelas, tidak ada jaminan bahwa jarak psikologis antara sangat buruk dan buruk sama dengan jarak antara baik dan sangat baik (Vetter, 2017).

Namun dalam praktik, skor total atau skor subskala dari kuesioner multi-item seperti ini seringkali dijumlahkan dan diperlakukan sebagai data numerik (interval). Justifikasinya adalah bahwa dengan menjumlahkan banyak item ordinal, skor total yang dihasilkan mulai mendekati sifat skala interval dan seringkali memiliki distribusi yang mendekati normal, terutama dengan ukuran sampel yang besar. Keputusan ini memiliki implikasi penting untuk analisis (Kim et al., 2024). Jika skor QoL diperlakukan sebagai data ordinal, uji non-parametrik seperti Mann-Whitney U akan digunakan untuk membandingkan kedua kelompok. Jika diperlakukan sebagai data numerik, *Independent T-Test* dapat digunakan, asalkan asumsi normalitas terpenuhi. Keputusan ini harus dinyatakan dan dibenarkan dengan jelas dalam protokol penelitian.

1.5.3 Simulasi Pemilihan *Flowchart* Uji pada Protokol Penelitian

Mari kita simulasikan penggunaan *flowchart* untuk memilih uji pada hipotesis utama studi *CardioNorm*.

- **Langkah 1: Apa tujuan utama analisis?** Tujuannya adalah membandingkan hasil (penurunan TDS) antarkelompok. Jadi, kita masuk ke cabang Komparasi/Perbedaan.
- **Langkah 2: Berapa banyak kelompok yang dibandingkan?** Ada dua kelompok: *CardioNorm* dan plasebo. Kita mengikuti cabang Dua Kelompok.
- **Langkah 3: Apakah kelompok-kelompok tersebut independen atau berpasangan?** Kelompok ini terdiri atas partisipan yang berbeda sehingga mereka independen. Kita mengikuti cabang Independen.
- **Langkah 4: Apa skala pengukuran variabel dependen (hasil)?** Variabel hasilnya adalah perubahan TDS yang merupakan data numerik (rasio). Kita mengikuti cabang Numerik/Kuantitatif.
- **Langkah 5 (Titik Keputusan Akhir):** *Flowchart* kini mengarah pada sebuah kotak yang menyajikan pilihan berdasarkan distribusi data. Jika data terdistribusi normal, gunakan

Independent T-Test. atau Jika data tidak terdistribusi normal, gunakan **Mann-Whitney U Test**. (Venkatesh & Santhosh, 2025).

Simulasi ini menunjukkan bagaimana *flowchart* secara efisien dan logis memandu peneliti dari pertanyaan penelitian yang luas ke uji statistik spesifik yang paling mungkin sesuai. Protokol penelitian akan menyatakan bahwa analisis primer akan menggunakan *Independent T-Test*, dengan ketentuan bahwa asumsi akan diperiksa terlebih dahulu dan Uji Mann-Whitney U akan digunakan sebagai analisis alternatif jika asumsi normalitas dilanggar secara signifikan.

Rangkuman Bab

- Statistika dalam riset Kedokteran berfungsi sebagai alat untuk merancang penelitian, menganalisis data, dan menginterpretasikan hasil sebagai landasan EBM.
- Statistik dibagi menjadi dua jenis utama: statistik deskriptif (meringkas data sampel) dan statistik inferensial (membuat generalisasi dari sampel ke populasi).
- Integritas data dan bioetika adalah prasyarat fundamental yang menjamin validitas dan tanggung jawab moral dalam penelitian.
- Data diklasifikasikan ke dalam empat skala pengukuran: nominal, ordinal, interval, dan rasio. Dua yang pertama adalah data kategorik, dan dua yang terakhir adalah data numerik.
- Pemilihan uji statistik mengikuti alur logika yang sistematis, dimulai dari perumusan pertanyaan penelitian dan hipotesis, penentuan tujuan uji, identifikasi jenis data, dan pemeriksaan asumsi.
- Validitas eksternal sebuah penelitian sangat bergantung pada proses *sampling* yang benar, yaitu pemilihan sampel yang representatif dari populasi target, dan penentuan ukuran sampel yang memadai.

- Kriteria inklusi dan eksklusi digunakan untuk mendefinisikan populasi studi secara spesifik, memastikan keamanan, dan meningkatkan homogenitas sampel.

Latihan Mahasiswa

Soal Esai

1. Jelaskan dengan kata-kata Anda sendiri perbedaan mendasar antara statistik deskriptif dan inferensial. Berikan contoh spesifik dari sebuah penelitian klinis fiktif untuk masing-masing jenis statistik tersebut!
2. Mengapa pemahaman tentang skala pengukuran data (nominal, ordinal, interval, rasio) dianggap sebagai langkah pertama yang paling krusial dalam memilih uji statistik? Jelaskan konsekuensi dari kesalahan identifikasi skala data!
3. Uraikan alur berpikir logis yang akan Anda gunakan untuk memilih uji statistik, mulai dari saat Anda memiliki ide penelitian hingga sebelum melakukan analisis. Sebutkan setidaknya empat langkah atau pertanyaan kunci dalam proses tersebut.
4. Jelaskan mengapa proses randomisasi (baik *random sampling* maupun *random assignment*) dianggap sangat penting dalam penelitian kedokteran. Apa saja potensi masalah yang dapat timbul jika randomisasi tidak dilakukan dengan benar?
5. Seorang peneliti menemukan bahwa asumsi normalitas data untuk uji-t tidak terpenuhi. Apa pilihan yang tersedia bagi peneliti tersebut? Jelaskan kelebihan dan kekurangan dari masing-masing pilihan.

Soal Pilihan Ganda

1. Seorang peneliti ingin meringkas karakteristik demografis partisipan penelitiannya, termasuk usia rata-rata dan persentase jenis kelamin. Jenis statistik apa yang paling sesuai untuk tujuan ini?
 - A. Statistik Inferensial,
 - B. Statistik Parametrik,

- C. Statistik Deskriptif -V-,
D. Statistik Non-Parametrik.
2. Variabel stadium kanker yang diklasifikasikan sebagai stadium I, II, III, dan IV adalah contoh dari skala pengukuran data...
A. Nominal,
B. Ordinal -V-,
C. Interval,
D. Rasio.
3. Seorang peneliti ingin membandingkan efektivitas tiga jenis obat penurun kolesterol. Variabel dependennya adalah kadar kolesterol LDL (data numerik). Ada tiga kelompok independen. Jika data terdistribusi normal, uji apakah yang paling tepat?
A. Uji-t Independen,
B. Uji *Chi-Square*,
C. ANOVA -V-,
D. Korelasi Pearson.
4. Apa tujuan utama dari penetapan kriteria eksklusi dalam sebuah protokol penelitian klinis?
A. Untuk memastikan sampel semirip mungkin dengan populasi umum,
B. Untuk meningkatkan jumlah partisipan yang dapat direkrut,
C. Untuk memperluas generalisasi hasil penelitian,
D. Untuk mengeluarkan individu yang mungkin berisiko atau dapat mengacaukan hasil. -V-
5. Hipotesis nol (H_0) dalam pengujian statistik biasanya merupakan pernyataan tentang...
A. Adanya efek atau perbedaan yang signifikan,
B. Tidak adanya efek atau tidak adanya perbedaan -V-,
C. Hubungan sebab-akibat antar variabel,
D. Kepastian 100% tentang parameter populasi.
6. Suatu studi yang memiliki ukuran sampel terlalu kecil kemungkinan besar akan mengalami...
A. Peningkatan risiko kesalahan Tipe I (positif palsu),
B. Kekuatan statistik yang rendah dan peningkatan risiko kesalahan Tipe II (negatif palsu) -V-,

- C. Validitas internal yang lebih tinggi,
 - D. Bias seleksi yang lebih sedikit.
7. Uji Mann-Whitney U adalah alternatif non-parametrik untuk...
- A. Uji-t Berpasangan,
 - B. ANOVA,
 - C. Uji-t Independen -V-,
 - D. Korelasi Pearson.
8. Suhu tubuh yang diukur dalam derajat Celcius adalah contoh data dengan skala...
- A. Nominal,
 - B. Ordinal,
 - C. Interval -V-,
 - D. Rasio.
9. Dalam kerangka PICO, C merupakan singkatan dari...
- A. *Conclusion*,
 - B. *Comparison* -V-,
 - C. *Criteria*,
 - D. *Confounding*.
10. Proses menerapkan fungsi matematika pada data untuk mengubah distribusinya agar memenuhi asumsi uji statistik disebut...
- A. Transformasi Variabel -V-,
 - B. Standardisasi,
 - C. Imputasi Data,
 - D. Analisis Sensitivitas.

Studi Kasus atau Tugas Kontekstual

1. Anda diminta untuk merancang sebuah studi untuk mengevaluasi efektivitas program edukasi gizi baru pada pengetahuan gizi (diukur dengan skor kuesioner dari 0-100) dan status gizi (diukur sebagai IMT, kg/m²) pada siswa sekolah dasar. Terdapat dua kelompok: kelompok yang menerima edukasi dan kelompok kontrol. Pengukuran dilakukan sebelum dan sesudah program.

- a. Rumuskan pertanyaan penelitian utama menggunakan kerangka PICO!
- b. Identifikasi variabel independen dan dependen utama dalam studi ini.
- c. Klasifikasikan setiap variabel berdasarkan skala pengukurannya.
- d. Untuk membandingkan perubahan skor pengetahuan antara kedua kelompok, uji statistik apa yang akan menjadi kandidat utama Anda? Jelaskan alur logika Anda.

Glosarium Bab

- **Bioetika:** Studi tentang isu-isu etis yang muncul dari kemajuan dalam bidang biologi dan kedokteran.
- **Data Kategorik:** Data yang nilainya mewakili kelompok atau kategori, seperti jenis kelamin atau stadium penyakit. Terdiri atas skala nominal dan ordinal.
- **Data Numerik:** Data yang nilainya mewakili kuantitas terukur, seperti tinggi badan atau tekanan darah. Terdiri atas skala interval dan rasio.
- **Hipotesis Nol (H_0):** Pernyataan statistik yang mengasumsikan tidak ada efek, tidak ada perbedaan, atau tidak ada hubungan antara variabel yang sedang diteliti.
- **Populasi:** Keseluruhan kelompok individu atau objek yang menjadi fokus penelitian dan dari mana kesimpulan akan ditarik.
- **Reliabilitas:** Tingkat konsistensi atau stabilitas suatu alat ukur dalam menghasilkan hasil yang sama pada pengukuran berulang.
- **Sampel:** Subkelompok atau bagian dari populasi yang dipilih untuk berpartisipasi dalam penelitian.
- **Sampling:** Proses pemilihan sampel dari populasi.
- **Statistik Deskriptif:** Cabang statistika yang berkaitan dengan peringkasan dan penyajian data.

- **Statistik Inferensial:** Cabang statistika yang menggunakan data sampel untuk membuat kesimpulan atau generalisasi tentang populasi.
- **Validitas:** Tingkat ketepatan suatu alat ukur dalam mengukur apa yang seharusnya diukur.
- **Variabel:** Karakteristik atau atribut yang dapat memiliki nilai yang berbeda untuk individu atau objek yang berbeda.

Daftar Pustaka Bab

- Al-Jundi, A., & Sakka, S. (2022). Evidence-Based Medicine: A Core Concept in Clinical Practice. *Cureus*, 14(3), e23447. <https://doi.org/10.7759/cureus.23447>
- Berger, V., Bour, L., Carter, K., Chipman, J., Everett, C., Heussen, N., Hewitt, C., Hilgers, R., Luo, Y., Renteria, J., Ryznik, Y., Sverdlov, O., Uschner, D., & Beckman, R. (2021). A roadmap to using randomization in clinical trials. *BMC Medical Research Methodology*, 21. <https://doi.org/10.1186/s12874-021-01303-z>
- Chicco, D., Sichenze, A., & Jurman, G. (2025). A simple guide to the use of Student's t-test, Mann-Whitney U test, Chi-squared test, and Kruskal-Wallis test in biostatistics. *BioData Mining*, 18. <https://doi.org/10.1186/s13040-025-00465-6>
- Christensen, R., Ranstam, J., Overgaard, S., & Wagner, P. (2023). Guidelines for a structured manuscript: Statistical methods and reporting in biomedical research journals. *Acta Orthopaedica*, 94, 243-249. <https://doi.org/10.2340/17453674.2023.11656>
- Guzik, P., & Więckowska, B. (2023). Data distribution analysis – a preliminary approach to quantitative data in biomedical research. *Journal of Medical Science*. <https://doi.org/10.20883/medical.e869>
- Kim, J., Kim, D., & Kwak, S. (2024). Comprehensive guidelines for appropriate statistical analysis methods in research. *Korean Journal of Anesthesiology*, 77, 503-517. <https://doi.org/10.4097/kja.24016>

- MyGPNNotes. Statistics. How to read a forest plot. 13, 2024. <https://www.mygpnnotes.com/statistics/how-to-read-a-forest-plot/>(Accessed on Januari 27, 2026)
- Pugh, S., & Torres-Saavedra, P. (2021). Fundamental Statistical Concepts in Clinical Trials and Diagnostic Testing. *The Journal of Nuclear Medicine*, 62, 757-764. <https://doi.org/10.2967/jnumed.120.245654>
- Resnik, D. B. (2023). Data management and sharing: Ethical issues for researchers. *Journal of Empirical Research on Human Research Ethics*, 18(4), 209-218. <https://doi.org/10.1177/15562646231191028>
- Serdar, C. C., Cihan, M., Yücel, D., & Serdar, M. A. (2021). Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. *Biochemia Medica*, 31(1), 010502. <https://doi.org/10.11613/BM.2021.010502>
- Tenny, S., & Abdelgawad, I. (2023). *Medical Statistics*. In StatPearls. StatPearls Publishing.
- Thabane, L., et al. (2021). A tutorial on the development of a research question, study objectives and hypotheses in clinical research. *Canadian Journal of Anesthesia/Journal canadien d'anesthésie*, 68(5), 711-721. <https://doi.org/10.1007/s12630-021-01921-5>
- Venkatesh, U., & Santhosh, V. (2025). Aligning Research Questions with Statistical Tests: A Clinician's Practical Framework. *NMO Journal*. https://doi.org/10.4103/jnmo.jnmo_61_25
- Vetter, T. R. (2017). Fundamentals of Research Data and Variables: The Devil Is in the Details. *Anesthesia & Analgesia*, 125(4), 1375–1380. <https://doi.org/10.1213/ANE.0000000000002374>

- **Odds Ratio (OR):** Ukuran asosiasi antara paparan dan hasil. Rasio *odds* hasil pada kelompok terpapar terhadap *odds* hasil pada kelompok tidak terpapar.
- **Output Viewer.** Jendela dalam perangkat lunak statistik (seperti SPSS) yang menampilkan semua hasil analisis, tabel, dan grafik.
- **Outlier (Pencilan):** Sebuah titik data yang terletak jauh dari titik-titik data lainnya dalam sampel.
- **P-hacking.** Praktik analisis data yang fleksibel yang meningkatkan kemungkinan menemukan hasil positif palsu, seperti dengan hanya melaporkan analisis yang signifikan secara statistik.
- **Pelabelan (Labelling):** Proses melampirkan label deskriptif ke nama variabel dan nilai numerik yang dikodekan.
- **Pengkodean (Coding):** Proses menetapkan nilai numerik ke respons kategorik.
- **Perbedaan Rerata (Mean Difference):** Ukuran efek absolut yang dihitung dengan mengurangi rerata satu kelompok dari rerata kelompok lainnya.
- **Peringkat (Rank):** Nilai posisi ordinal dari sebuah observasi dalam serangkaian data yang diurutkan.
- **Plot QQ (Quantile-Quantile):** Perangkat grafis yang digunakan untuk membandingkan dua distribusi probabilitas dengan memplot kuantil mereka satu sama lain.
- **Populasi:** Keseluruhan kelompok individu atau objek yang menjadi fokus penelitian dan dari mana kesimpulan akan ditarik.
- **Proportional Hazard Assumption.** Asumsi inti dari model Cox yang menyatakan bahwa rasio *hazard* untuk setiap dua individu adalah konstan dari waktu ke waktu.
- **Rasio Likelihood (Likelihood Ratio - LR):** Ukuran kinerja tes diagnostik yang menunjukkan seberapa besar hasil tes mengubah *odds* penyakit.

- **Reduksi Dimensi:** Proses mengurangi jumlah variabel acak yang dipertimbangkan, melalui pemilihan variabel atau ekstraksi fitur.
- **Regresi Linier Berganda:** Teknik regresi yang menggunakan dua atau lebih variabel independen untuk memprediksi satu variabel dependen numerik.
- **Regresi Linier Sederhana:** Teknik regresi yang menggunakan satu variabel independen untuk memprediksi satu variabel dependen numerik.
- **Regresi Logistik Biner:** Model regresi yang digunakan untuk memprediksi hasil dari variabel dependen yang hanya memiliki dua kategori.
- **Regresi Logistik Multinomial:** Model regresi yang digunakan untuk memprediksi hasil dari variabel dependen nominal dengan lebih dari dua kategori.
- **Regresi Logistik Ordinal:** Model regresi yang digunakan untuk memprediksi hasil dari variabel dependen yang kategorinya memiliki urutan.
- **Relative Risk (RR):** Rasio probabilitas suatu hasil pada kelompok terpapar terhadap probabilitas hasil pada kelompok tidak terpapar. Juga dikenal sebagai Rasio Risiko.
- **Reliabilitas:** Konsistensi atau stabilitas dari suatu alat ukur.
- **Rentang Interkuartil (*Interquartile Range* - IQR):** Ukuran penyebaran statistik yang dihitung sebagai selisih antara kuartil ketiga (Q3) dan kuartil pertama (Q1).
- **Repeated Measures ANOVA:** Jenis ANOVA yang digunakan ketika variabel yang sama diukur pada subjek yang sama di bawah tiga atau lebih kondisi berbeda.
- **Rerata yang Disesuaikan (*Adjusted Mean*):** Rerata kelompok yang diestimasi oleh model ANCOVA, yang mewakili rerata yang diharapkan jika semua kelompok memiliki nilai rata-rata yang sama pada kovariat.
- **Residual:** Selisih antara nilai dependen yang diamati secara aktual dan nilai yang diprediksi oleh model regresi ($Y_{\text{aktual}} - Y_{\text{prediksi}}$).

- **Robustness:** Sifat dari prosedur statistik yang kinerjanya tidak terlalu terpengaruh oleh pelanggaran asumsi.
- **Sampel:** Subkelompok atau bagian dari populasi yang dipilih untuk berpartisipasi dalam penelitian.
- **Sampling:** Proses pemilihan sampel dari populasi.
- **Scatter Plot (Diagram Pencar):** Grafik yang digunakan untuk memvisualisasikan hubungan antara dua variabel numerik.
- **Sensitivitas:** Proporsi individu dengan penyakit yang memiliki hasil tes positif.
- **Signifikansi Klinis:** Relevansi atau pentingnya temuan penelitian dalam praktik klinis; apakah besaran efek cukup besar untuk membuat perbedaan bagi pasien.
- **Signifikansi Statistik:** Hasil yang tidak mungkin terjadi karena kebetulan, biasanya didefinisikan oleh *p-value* di bawah ambang batas (misalnya, $p < 0,05$).
- **Skor Selisih (Difference Score):** Nilai yang dihitung untuk setiap pasangan dalam data dependen, biasanya dengan mengurangi pengukuran pertama dari pengukuran kedua.
- **Spesifisitas:** Proporsi individu tanpa penyakit yang memiliki hasil tes negatif.
- **Sphericity:** Asumsi dari *Repeated Measures* ANOVA bahwa varians dari perbedaan antara semua kemungkinan pasangan kondisi adalah sama.
- **Statistik Deskriptif:** Cabang statistika yang berkaitan dengan peringkasan dan penyajian data.
- **Statistik Inferensial:** Cabang statistika yang menggunakan data sampel untuk membuat kesimpulan atau generalisasi tentang populasi.
- **Statistik Non-Parametrik:** Metode statistik yang tidak membuat asumsi tentang distribusi probabilitas dari populasi yang dianalisis.
- **Statistik Parametrik:** Metode statistik yang mengasumsikan bahwa data berasal dari jenis distribusi probabilitas tertentu (misalnya, normal) dan membuat inferensi tentang parameter dari distribusi tersebut.

- ***Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)***: Pedoman pelaporan untuk studi observasional.
- **Tabel Kontingensi**: Tabel yang menampilkan distribusi frekuensi dari variabel-variabel. Juga dikenal sebagai tabel silang.
- **Teorema Batas Sentral (*Central Limit Theorem*)**: Teorema fundamental yang menyatakan bahwa distribusi sampling dari rerata sampel akan mendekati distribusi normal seiring meningkatnya ukuran sampel.
- **Transformasi Data**: Penerapan fungsi matematika ke setiap titik data untuk mengatasi masalah seperti non-normalitas atau heterogenitas varians.
- **Transformasi Variabel**: Proses membuat variabel baru dengan menerapkan fungsi matematika ke variabel yang ada.
- **Uji *Chi-Square* (χ^2) Pearson**: Uji statistik yang digunakan untuk menentukan apakah ada asosiasi yang signifikan antara dua variabel kategorik.
- **Uji Cochran's Q**: Uji non-parametrik untuk membandingkan lebih dari dua kelompok berpasangan atau berulang di mana variabel hasilnya adalah biner.
- **Uji Dunn**: Uji *post-hoc* non-parametrik yang sering digunakan setelah Uji Kruskal-Wallis yang signifikan.
- **Uji Eksak Fisher (*Fisher's Exact Test*)**: Uji statistik yang digunakan untuk analisis tabel kontingensi (biasanya 2x2) ketika ukuran sampel kecil.
- **Uji Friedman**: Uji non-parametrik untuk menganalisis data dari tiga atau lebih pengukuran berulang atau kondisi berpasangan.
- **Uji Hosmer-Lemeshow**: Uji statistik untuk menilai kelayakan atau kalibrasi dari sebuah model regresi logistik.
- **Uji Kolmogorov-Smirnov**: Uji statistik untuk menentukan apakah suatu sampel berasal dari distribusi tertentu (misalnya, distribusi normal).

- **Uji Kruskal-Wallis:** Uji non-parametrik untuk membandingkan tiga atau lebih kelompok independen. Alternatif untuk *One-Way* ANOVA.
- **Uji Levene:** Uji statistik untuk menilai kesamaan varians antara dua atau lebih kelompok.
- **Uji Log-Rank:** Uji hipotesis non-parametrik untuk membandingkan dua atau lebih fungsi ketahanan hidup.
- **Uji Mann-Whitney U:** Uji non-parametrik untuk membandingkan dua sampel independen. Kadang disebut Uji Peringkat-Jumlah Wilcoxon.
- **Uji Mauchly:** Uji statistik yang digunakan untuk menguji asumsi *sphericity*.
- **Uji McNemar:** Uji non-parametrik untuk menganalisis data kategorik biner dari desain berpasangan (pra-pasca).
- **Uji Post-Hoc:** Uji statistik yang dilakukan setelah uji omnibus (seperti ANOVA) yang signifikan untuk menentukan perbedaan spesifik antar pasangan kelompok.
- **Uji Shapiro-Wilk:** Uji statistik dengan kekuatan tinggi untuk menguji hipotesis nol bahwa suatu sampel berasal dari populasi berdistribusi normal.
- **Uji-t Berpasangan (Paired T-Test):** Uji statistik untuk membandingkan rerata dari dua sampel terkait atau dependen.
- **Uji-t Independen (Independent T-Test):** Uji statistik untuk membandingkan rerata dari dua sampel independen.
- **Ukuran Efek (Effect Size):** Ukuran kuantitatif dari kekuatan atau besarnya suatu fenomena atau hubungan.
- **Validitas:** Sejauh mana sebuah instrumen mengukur apa yang seharusnya diukur.
- **Validitas Isi (Content Validity):** Sejauh mana item-item dalam suatu instrumen mewakili semua aspek dari konstruk yang diukur.
- **Validitas Konstruk (Construct Validity):** Sejauh mana suatu instrumen secara akurat mengukur konstruk teoretis yang mendasarinya.

- **Variabel:** Karakteristik atau atribut yang dapat memiliki nilai yang berbeda untuk individu atau objek yang berbeda.
- **Variabel Pengganggu (*Confounding Variable*):** Variabel ketiga yang berhubungan dengan variabel independen dan dependen, yang dapat mendistorsi hubungan yang sebenarnya di antara keduanya.
- **Waktu-ke-Kejadian (*Time-to-Event*):** Durasi waktu dari titik awal yang ditentukan hingga terjadinya suatu peristiwa yang diminati.
- **Welch's t-test:** Adaptasi dari Uji-t Student yang digunakan ketika dua sampel memiliki varians yang tidak sama.
- **Wilks' Lambda:** Salah satu statistik uji utama yang digunakan dalam MANOVA untuk menilai signifikansi perbedaan antar kelompok.



www.penerbitbukumurah.com

Dilarang keras, mencetak naskah
hasil layout ini tanpa seijin Penerbit



www.penerbitbukumurah.com

**Dilarang keras, mencetak naskah
hasil layout ini tanpa seijin Penerbit**



REFERENSI

- Aggarwal, C. C. (2017). *Outlier Analysis* (2nd ed.). Springer.
- Akobeng, A. K. (2007). Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta paediatrica (Oslo, Norway : 1992)*, 96(5), 644–647. <https://doi.org/10.1111/j.1651-2227.2006.00178.x>
- Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish journal of emergency medicine*, 18(3), 91–93. <https://doi.org/10.1016/j.tjem.2018.08.001>
- Al-Jundi, A., & Sakka, S. (2022). Evidence-Based Medicine: A Core Concept in Clinical Practice. *Cureus*, 14(3), e23447. <https://doi.org/10.7759/cureus.23447>
- Andrade, C. (2015). Understanding the relative risk, odds ratio, and related terms: as simple as it can be. *The Journal of clinical psychiatry*, 76(7), e857–e861. <https://doi.org/10.4088/JCP.15f10150>
- Asuero, A. G., Sayago, A., & González, G. (2006). The correlation coefficient: An overview. *Critical reviews in analytical chemistry*, 36(1), 41–59.
- Bender, R., & Grouven, U. (1998). Using binary logistic regression to analyse ordinal data. *Statistical methods in medical research*, 7(2), 157–168.
- Berger, V., Bour, L., Carter, K., Chipman, J., Everett, C., Heussen, N., Hewitt, C., Hilgers, R., Luo, Y., Renteria, J., Ryznik, Y., Sverdlov, O., Uschner, D., & Beckman, R. (2021). A roadmap to using randomization in clinical trials. *BMC Medical Research Methodology*, 21. <https://doi.org/10.1186/s12874-021-01303-z>

- Bland, J. M., & Altman, D. G. (2004). The logrank test. *BMJ (Clinical research ed.)*, 328(7447), 1073. <https://doi.org/10.1136/bmj.328.7447.1073>
- Bland, M. (2023). *An Introduction to Medical Statistics* (5th ed.). Oxford University Press.
- Bonett, D. G., & Wright, T. A. (2000). Sample size requirements for estimating Pearson, Kendall and Spearman correlations. *Psychometrika*, 65(1), 23-28.
- Card, A., & pendiente, K. (2023). A Tutorial on the Paired-Samples t-test. *Scholarship of Teaching and Learning*, 6(2), 35-51.
- Casson, R. J., & Farmer, L. D. (2014). Understanding and checking the assumptions of linear regression: a primer for medical researchers. *Clinical & experimental ophthalmology*, 42(6), 590–596. <https://doi.org/10.1111/ceo.12321>
- Chicco, D., Sichenze, A., & Jurman, G. (2025). A simple guide to the use of Student's t-test, Mann-Whitney U test, Chi-squared test, and Kruskal-Wallis test in biostatistics. *BioData Mining*, 18. <https://doi.org/10.1186/s13040-025-00465-6>
- Christensen, R., Ranstam, J., Overgaard, S., & Wagner, P. (2023). Guidelines for a structured manuscript: Statistical methods and reporting in biomedical research journals. *Acta Orthopaedica*, 94, 243-249. <https://doi.org/10.2340/17453674.2023.11656>
- Clark, T. G., Bradburn, M. J., Love, S. B., & Altman, D. G. (2003). Survival analysis part I: basic concepts and first analyses. *British journal of cancer*, 89(2), 232–238. <https://doi.org/10.1038/sj.bjc.6601118>
- Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for psychometric instruments: theory and application. *The American journal of medicine*, 119(2), 166.e7–166.e16. <https://doi.org/10.1016/j.amjmed.2005.10.036>
- Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2), 187-202.
- Das, K. R., & Imon, A. H. M. R. (2016). A brief review of tests for normality. *American Journal of Theoretical and Applied Statistics*, 5(1), 5-12. <https://doi.org/10.11648/j.ajtas.20160501.12>

- De Oliveira Santana Amaral, E., & Line, S. (2021). Current use of effect size or confidence interval analyses in clinical and biomedical research. *Scientometrics*, 126, 9133 - 9145. <https://doi.org/10.1007/s11192-021-04150-3>
- Deeks, J. J., & Altman, D. G. (2004). Diagnostic tests 4: likelihood ratios. *BMJ (Clinical research ed.)*, 329(7458), 168–169. <https://doi.org/10.1136/bmj.329.7458.168>
- Derrick, B., Toher, D., & White, P. (2016). Why Welch's test is Type I error robust. *The Quantitative Methods for Psychology*, 12(1), 30-38. <https://doi.org/10.20982/tqmp.12.1.p030>
- Dinno, A. (2015). Nonparametric pairwise multiple comparisons in independent groups using Dunn's test. *The Stata Journal*, 15(1), 292-300.
- Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis* (5th ed.). John Wiley & Sons.
- Fagerland, M. W., & Hosmer, D. W. (2012). A goodness-of-fit test for the proportional odds logistic regression model. *Statistics in medicine*, 31(8), 819–834. <https://doi.org/10.1002/sim.4463>
- Faltin, F. W. (2022). *Statistical Power and Sample Size in Clinical Trials*. Springer.
- Feller, W. (2020). The fundamental limit theorems in probability. *Bulletin of the American Mathematical Society*, 51(12), 800-832. (Reprint).
- Ferreira, J. C., & Patino, C. M. (2015). Understanding clinical significance, statistical significance, and the presentation of quantitative results. *Jornal Brasileiro de Pneumologia*, 41(5), 397. <https://doi.org/10.1590/S1806-37132015000000212>
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Sage publications.
- Fullerton, A. S. (2009). A conceptual framework for ordered logistic regression models. *Sociological Methods & Research*, 38(2), 306-347.
- Gallagher, C., O'Connell, T., & O'Sullivan, L. (2019). A practical guide to the analysis of covariance (ANCOVA) in biomedical research. *Journal of Clinical Investigation*, 129(5), 1845-1847.

- Gauthier, T. D. (2001). Detecting trends using Spearman's rank correlation coefficient. *Environmental Forensics*, 2(4), 359-362.
- Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *International journal of endocrinology and metabolism*, 10(2), 486-489. <https://doi.org/10.5812/ijem.3505>
- Goligher, E., Heath, A., & Harhay, M. (2024). Bayesian statistics for clinical research. *The Lancet*, 404, 1067-1076. [https://doi.org/10.1016/s0140-6736\(24\)01295-9](https://doi.org/10.1016/s0140-6736(24)01295-9)
- Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 81(3), 515-526.
- Gueorguieva, R., & Krystal, J. H. (2004). Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. *Archives of general psychiatry*, 61(3), 310-317. <https://doi.org/10.1001/archpsyc.61.3.310>
- Guzik, P., & Więckowska, B. (2023). Data distribution analysis – a preliminary approach to quantitative data in biomedical research. *Journal of Medical Science*. <https://doi.org/10.20883/medical.e869>
- Hajian-Tilaki, K. (2013). Receiver Operating Characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian journal of internal medicine*, 4(2), 627-635.
- Hauke, J., & Kossowski, T. (2011). Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones geographicae*, 30(2), 87-93.
- Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS biology*, 13(3), e1002106. <https://doi.org/10.1371/journal.pbio.1002106>
- Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-based nursing*, 18(3), 66-67. <https://doi.org/10.1136/eb-2015-102129>

- Huberty, C. J., & Olejnik, S. (2006). *Applied MANOVA and discriminant analysis* (2nd ed.). John Wiley & Sons.
- Hughes, R. A., Tilling, K., & Sterne, J. A. C. (2021). Multiple imputation for handling missing data in medical research: a practical guide with examples. *Journal of clinical epidemiology*, 138, 169-180.
- Johnson, M. (2020). *Data Management and Coding for Social and Behavioral Sciences*. SAGE Publications.
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202. <https://doi.org/10.1098/rsta.2015.0202>
- Kamath, A., Poojari, S., & Varsha, K. (2025). Assessing the robustness of normality tests under varying skewness and kurtosis: a practical checklist for public health researchers. *BMC Medical Research Methodology*, 25. <https://doi.org/10.1186/s12874-025-02641-y>
- Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282), 457-481.
- Kasuya, E. (2023). Type I and type II errors in statistical testing. *Ecological Research*, 38(2), 171-175. <https://doi.org/10.1111/1440-1703.12357>
- Kim, H. Y. (2017). Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. *Restorative dentistry & endodontics*, 42(2), 152-155. <https://doi.org/10.5395/rde.2017.42.2.152>
- Kim, H. Y. (2017). Statistical notes for clinical researchers: Post-hoc analysis. *Restorative dentistry & endodontics*, 42(3), 231-235. <https://doi.org/10.5395/rde.2017.42.3.231>
- Kim, J., Kim, D., & Kwak, S. (2024). Comprehensive guidelines for appropriate statistical analysis methods in research. *Korean Journal of Anesthesiology*, 77, 503-517. <https://doi.org/10.4097/kja.24016>
- Kim, S. (2015). A guide to the appropriate use of correlation coefficient in medical research. *Korean Journal of Anesthesiology*, 68(3), 299-300. <https://doi.org/10.4097/kjae.2015.68.3.299>

- Kim, T. K., & Park, J. H. (2019). More about the basic assumptions of t-test: normality and sample size. *Korean journal of anesthesiology*, 72(4), 331–335. <https://doi.org/10.4097/kja.19018>
- Klein, J. P., & Moeschberger, M. L. (2003). *Survival analysis: techniques for censored and truncated data*. Springer.
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of chiropractic medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kumar, A., Kishun, J., Singh, U., Gaur, D., Mishra, P., & Pandey, C. (2023). Use of appropriate statistical tools in biomedical research: Current trend & status. *The Indian Journal of Medical Research*, 157, 353 - 357. https://doi.org/10.4103/ijmr.ijmr_809_20
- Kwack, S. Y., & Ko, K. (2021). The central limit theorem: The cornerstone of modern statistics. *Korean Journal of Anesthesiology*, 74(5), 442–443. <https://doi.org/10.4097/kja.21326>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology*, 4, 863. <https://doi.org/10.3389/fpsyg.2013.00863>
- Lakens, D. (2021). The practical alternative to significance testing. *BMC Psychology*, 9(1), 5. <https://doi.org/10.1186/s40359-021-00527-6>
- LaValley, M. P. (2022). Logistic regression. *Circulation*, 117(18), 2395–2399. <https://doi.org/10.1161/CIRCULATIONAHA.106.682658>
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel psychology*, 28(4), 563–575.
- Leys, C., Klein, O., Dominicy, Y., & Ley, C. (2019). Detecting multivariate outliers: Use a robust variant of the Mahalanobis distance. *Journal of Experimental Social Psychology*, 51, 34–39.
- Lix, L. M., & Sajobi, T. T. (2023). *Robust Statistical Methods for Health Sciences*. CRC Press.
- MacFarland, T. W., & Yates, J. M. (2016). *Introduction to nonparametric statistics for the biological sciences using R*. Springer.
- Madley-Dowd, P., Hughes, R., Tilling, K., & Heron, J. (2019). The proportion of missing data should not be used to guide decisions

- on multiple imputation. *Journal of clinical epidemiology*, 110, 63–73. <https://doi.org/10.1016/j.jclinepi.2019.02.016>
- Mansournia, M., & Nazemipour, M. (2024). Recommendations for accurate reporting in medical research statistics. *The Lancet*, 403, 611–612. [https://doi.org/10.1016/s0140-6736\(24\)00139-9](https://doi.org/10.1016/s0140-6736(24)00139-9)
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3), 276–282.
- McHugh, M. L. (2013). The chi-square test of independence. *Biochemia medica*, 23(2), 143–149. <https://doi.org/10.11613/bm.2013.018>
- McHugh, M. L. (2015). The McNemar Test and Cochran's Q Test. *Biochemia medica*, 25(2), 173–191.
- McLachlan, G. J. (2004). *Discriminant analysis and statistical pattern recognition*. John Wiley & Sons.
- Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of cardiac anaesthesia*, 22(1), 67–72. https://doi.org/10.4103/aca.ACA_157_18
- Mukaka, M. M. (2012). A guide to appropriate use of Correlation coefficient in medical research. *Malawi medical journal*, 24(3), 69–71.
- MyGPNNotes. Statistics. How to read a forest plot. 13, 2024. <https://www.mygpnnotes.com/statistics/how-to-read-a-forest-plot/> (Accessed on Januari 27, 2026)
- Nahm, F. S. (2022). Nonparametric statistical tests for the continuous data: the basic concept and the practical use. *Korean Journal of Anesthesiology*, 75(1), 8–14. <https://doi.org/10.4097/kja.21287>
- O'Connell, A. A. (2022). A practical guide to log transformation of biomedical data. *Journal of Biopharmaceutical Statistics*, 32(4), 549–565.
- Oehring, D. (2025). Biostatistik in Ophthalmologie und Optometrie: Eine Artikelserie zur Unterstützung evidenzbasierter Entscheidungsprozesse Teil 2: Statistische Tests – Prinzipien und Anwendung. *Optometry & Contact Lenses*. <https://doi.org/10.54352/dozv.smzt8871>

- Ostertagová, E., Ostertag, O., & Kováč, J. (2023). Methodology and Application of the Two-Sample T-Test. *Applied Mechanics and Materials*, 475-476, 545-552. <https://doi.org/10.4028/www.scientific.net/AMM.475-476.545>
- Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (7th ed.). Routledge.
- Pannuti, C., Da Silva, H., Su, N., Li, K., & Heitz-Mayfield, L. (2025). Statistical Gems and Cautions From the Statistical Advisory Board of Clinical Oral Implants Research.. *Clinical oral implants research*. <https://doi.org/10.1111/clr.70033>
- Parikh, R., Mathai, A., Parikh, S., Chandra Sekhar, G., & Thomas, R. (2008). Understanding and using sensitivity, specificity and predictive values. *Indian journal of ophthalmology*, 56(1), 45–50. <https://doi.org/10.4103/0301-4738.37595>
- Park, S. H., & Goo, J. M. (2024). A practical guide to understanding and using the receiver operating characteristic (ROC) curve. *Korean Journal of Radiology*, 15(1), 11-18. <https://doi.org/10.3348/kjr.2014.15.1.11>
- Poole, M. A., & O'Farrell, P. N. (2022). The assumptions of the linear regression model. *Transactions of the Institute of British Geographers*, 52, 5-20. (Reprint).
- Pugh, S., & Torres-Saavedra, P. (2021). Fundamental Statistical Concepts in Clinical Trials and Diagnostic Testing. *The Journal of Nuclear Medicine*, 62, 757 - 764. <https://doi.org/10.2967/jnumed.120.245654>
- Ranstam, J. (2022). Why the P-value culture is bad and confidence intervals a better alternative. *Acta Orthopaedica*, 93, 1-2. <https://doi.org/10.2340/17453674.2022.2423>
- Resnik, D. B. (2023). Data management and sharing: Ethical issues for researchers. *Journal of Empirical Research on Human Research Ethics*, 18(4), 209-218. <https://doi.org/10.1177/15562646231191028>
- Rich, J. T., Neely, J. G., Paniello, R. C., Voelker, C. C., Nussenbaum, B., & Wang, E. W. (2010). A practical guide to understanding Kaplan-Meier curves. *Otolaryngology--head and neck surgery : official*

Journal of American Academy of Otolaryngology-Head and Neck Surgery, 143(3), 331–336. <https://doi.org/10.1016/j.otohns.2010.05.007>

- Rohrer, J. M. (2019). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, 1(1), 27-42.
- Rust, C. F., & Lauth, M. (2023). A comparison of tests for variance homogeneity. *PLOS ONE*, 18(9), e0291500. <https://doi.org/10.1371/journal.pone.0291500>
- Sawilowsky, S. S. (2021). Nonparametric tests of interaction in experimental design. *Journal of Modern Applied Statistical Methods*, 19(1), eP2862. <https://doi.org/10.22237/jmasm/1609459200>
- Schneider, A., Hommel, G., & Blettner, M. (2010). Linear regression analysis: part 14 of a series on evaluation of scientific publications. *Deutsches Arzteblatt international*, 107(44), 776–782. <https://doi.org/10.3238/arztebl.2010.0776>
- Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesthesia and analgesia*, 126(5), 1763–1768. <https://doi.org/10.1213/ANE.00000000000002864>
- Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ (Clinical research ed.)*, 340, c332. <https://doi.org/10.1136/bmj.c332>
- Serdar, C. C., Cihan, M., Yücel, D., & Serdar, M. A. (2021). Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. *Biochemia Medica*, 31(1), 010502. <https://doi.org/10.11613/BM.2021.010502>
- Sperandei, S. (2014). Understanding logistic regression analysis. *Biochemia medica*, 24(1), 12–18. <https://doi.org/10.11613/BM.2014.003>

- Spruance, S. L., Reid, J. E., Grace, M., & Samore, M. (2004). Hazard ratio in clinical trials. *Antimicrobial agents and chemotherapy*, 48(8), 2787–2792. <https://doi.org/10.1128/AAC.48.8.2787-2792.2004>
- Starkweather, J., & Moske, A. K. (2011). Multinomial logistic regression. *Journal of the American Statistical Association*, 106(494), 776–778.
- Sullivan, L. M. (2024). *Essentials of Biostatistics in Public Health* (4th ed.). Jones & Bartlett Learning.
- Szumilas, M. (2010). "Explaining odds ratios". *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 19(3), 227–229.
- Tabachnick, B. G., & Fidell, L. S. (2021). *Using multivariate statistics* (7th ed.). Pearson.
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273–1296.
- Taherdoost, H. (2023). *Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research*. In *Advanced Methodologies and Technologies in Business Operations and Management* (pp. 2043–2060). IGI Global.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Tenny, S., & Abdelgawad, I. (2023). *Medical Statistics*. In StatPearls. StatPearls Publishing.
- Thabane, L., et al. (2021). A tutorial on the development of a research question, study objectives and hypotheses in clinical research. *Canadian Journal of Anesthesia/Journal canadien d'anesthésie*, 68(5), 711–721. <https://doi.org/10.1007/s12630-021-01921-5>
- Trajman, A., & Luiz, R. R. (2014). McNemar χ^2 test: a sensible and useful alternative to the paired t test for the analysis of before-and-after studies. *Jornal Brasileiro de Pneumologia*, 40(4), III–IV. <https://doi.org/10.1590/S1806-37132014000400002>
- Turner, J. R., & Thayer, J. F. (2022). *Introduction to analysis of variance: Design, analysis, & interpretation*. Sage Publications.

- van den Broeck, J., Cunningham, S. A., Eeckels, R., & Herbst, K. (2005). Data cleaning: detecting, diagnosing, and editing data abnormalities. *PLoS medicine*, 2(10), e267. <https://doi.org/10.1371/journal.pmed.0020267>
- Venkatesh, U., & Santhosh, V. (2025). Aligning Research Questions with Statistical Tests: A Clinician's Practical Framework. *NMO Journal*. https://doi.org/10.4103/jnmo.jnmo_61_25
- Vetter, T. R. (2017). Fundamentals of Research Data and Variables: The Devil Is in the Details. *Anesthesia & Analgesia*, 125(4), 1375–1380. <https://doi.org/10.1213/ANE.0000000000002374>
- Vilakati, S. (2025). Prompt engineering for accurate statistical reasoning with large language models in medical research. *Frontiers in Artificial Intelligence*, 8. <https://doi.org/10.3389/frai.2025.1658316>
- Vittinghoff, E., Glidden, D. V., Shiboski, S. C., & McCulloch, C. E. (2012). *Regression methods in biostatistics: Linear, logistic, survival, and repeated measures models* (2nd ed.). Springer.
- von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet (London, England)*, 370(9596), 1453–1457. [https://doi.org/10.1016/S0140-6736\(07\)61602-X](https://doi.org/10.1016/S0140-6736(07)61602-X)
- Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology*, 44(3), 219-246.
- Wickham, H., & Grolemund, G. (2017). *R for data science: Import, tidy, transform, visualize, and model data*. O'Reilly Media, Inc.
- Winter, B. (2021). The interpretation of p-values and confidence intervals: A review for researchers. *Psychonomic Bulletin & Review*, 28(4), 1089-1100.
- Zimmerman, D. W., & Zumbo, B. D. (1993). Friedman's test: The proper alternative to the one-way repeated measures ANOVA for ranked data. *Perceptual and Motor Skills*, 77(2), 523-529. <https://doi.org/10.2466/pms.1993.77.2.523>

Zhou, Y., Zhu, Y., & Wong, W. (2023). Statistical tests for homogeneity of variance for clinical trials and recommendations. *Contemporary Clinical Trials Communications*, 33. <https://doi.org/10.1016/j.conctc.2023.101119>



www.penerbitbukumurah.com

Dilarang keras, mencetak naskah
hasil layout ini tanpa seijin Penerbit



PROFIL PENULIS



Prof. Dr. Drs. H. Kusmardi, M.Sc., CIRR adalah Guru Besar Tetap di Fakultas Kedokteran, Universitas Indonesia. Sekarang adalah mahasiswa Program Studi Doktor Ilmu Dakwah di Fakultas Ilmu Agama Islam, Universitas Islam Asy-Syafi'iyah Jakarta. Pendidikan Doktor sebelumnya ditempuh di Program Doktor Ilmu Biomedik FKUI Jakarta dan Program Doktor Ilmu Biomedis Hewan IPB Bogor. Program Magister di FKUI,

Program S1 di Biologi FMIPA UI. Pernah Belajar Biologi Molekular di Melbourne University. SD sampai SMA di kota kelahirannya, Sumenep, Madura. Bidang penelitian utamanya meliputi kanker kolorektal dan kanker payudara, dengan penekanan khusus pada eksplorasi potensi efek penghambatan berbagai obat alami Indonesia terhadap karsinogenesis pada kedua jenis kanker tersebut. Penelitiannya juga melibatkan identifikasi jaringan normal versus perkembangan kanker menggunakan penanda molekuler dan model komputasional. Selain itu, ia telah menulis beberapa buku, diantaranya: *The Mouse Model for Breast Cancer*; *The Mouse Model for Colorectal Cancer*; *Lunasin: A Soybean Polypeptide as a Chemopreventive Adjuvant for Colon Cancer*, dan Pembuatan Antibodi Poliklonal, Peristiwa Ilmiah yang telah Dijelaskan dalam Al-Qur'an, serta Manusia dan Semesta yang telah Dijelaskan dalam Al-Qur'an. Selama ini beliau sudah menghasilkan seratus lebih publikasi internasional bereputasi terindeks Scopus, puluhan hak kekayaan intelektual. Saat ini Prof. Kusmardi menjadi dosen *home base* di Program Studi Doktor Ilmu Biomedik FKUI, selain dosen tetap di Prodi Magister Ilmu Biomedik, Prodi Dokter Spesialis Patologi Anatomi, dan Prodi Pendidikan Dokter FKUI. Beliau secara

aktif menjadi *Reviewer* internal hibah UI dan *Reviewer* Nasional Proposal Hibah Kemendiktisaintek, Auditor Internal Akademik UI, PIC QA *Drug Development Research Center* IMERI FKUI, *Cancer drug discovery coordinator* di *Human Cancer Research Center* IMERI FKUI, tim penjamin mutu akademik Program Doktor Ilmu Kedokteran FKUI, serta sebagai Asesor LAM-PTKes divisi Kedokteran. Untuk mempertajam khasanah pengetahuannya, beliau juga menjadi Ketua Dewan Redaksi Majalah Patologi Indonesia, serta aktif menjadi reviewer tetap pada beberapa jurnal internasional terindeks scopus Q1. Kemampuan mengenalkan statistik diperoleh sejak mengajar statistic di prodi S1, Spesialis, dan S3 sejak lama.



Prof. Dr. Drs. H. Yurnadi, M.Kes. adalah Guru Besar Tetap di Fakultas Kedokteran Universitas Indonesia. Lahir di Agam 19 September 1966, anak ke 2 atas 5 bersaudara dari pasangan Bapak H. Hanafi Sutan Pamuntjak (Alm.) dan Ibu Hj. Martina Mukhtar (Almh.).

Menempuh Pendidikan dari SD, SMP, dan SMA di Kota Batu Sangkar Sumatera Barat. Menyelesaikan Pendidikan Sarjana di Jurusan Biologi Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Andalas pada tahun 1991. Kemudian mengabdikan sebagai Staf Pengajar di Departemen Biologi Kedokteran Fakultas Kedokteran Universitas Indonesia Jakarta sejak tahun 1994 sampai sekarang. Pada tahun 1996 penulis melanjutkan Pendidikan S2 di Program Studi Magister Ilmu Biomedik Fakultas Kedokteran Universitas Indonesia Jakarta atas beasiswa TMPD DIKTI dan mendapatkan gelar Magister Kesehatan (M.Kes.) pada tahun 1999. Pada tahun 2006, atas beasiswa FKUI dan BPPS DIKTI penulis melanjutkan Pendidikan S3 di Program Doktor Ilmu Biomedik Fakultas Kedokteran Universitas Indonesia Jakarta dan mendapatkan gelar Doktor (Dr.) pada tahun 2011.

Sejak tahun 1998 sampai sekarang penulis mengajar matakuliah Embriologi, Biologi Reproduksi, Genetika Kedokteran, Andrologi,

Embriologi Molekuler, Modul Neurosain dan Modul Tumbuh Kembang, Matakuliah Pengembangan Kepribadian dan Terintegrasi (MPKT) pada Program Pendidikan S1 FKUI dan Pendidikan Magister Ilmu Biomedik FKUI. Di samping itu, penulis juga pernah mengajar matakuliah MPKT di Tingkat Universitas Indonesia, baik di FKG, FISIP, FKM dan FT.

Selain sebagai Staf Pengajar, penulis juga pernah menjabat sebagai Koordinator Administrasi dan Keuangan Departemen Biologi Kedokteran pada beberapa periode sejak 2001-2021, menjabat sebagai Ketua TIM PAK FKUI 2012-2014, Ketua Modul MPKT FKUI 2014-2016, Tim Pengembang Modul MPKT UI 2015-2016, Anggota Senat Akademik FKUI pada Komisi 4 dan Komisi 1 periode 2018 sampai sekarang, Anggota Dewan Guru Besar FKUI di Komite 1 dan Komite 5 sejak tahun 2023 sampai sekarang.

Penulis juga banyak melakukan penelitian untuk mengeksplorasi efektivitas bahan alam (Biji Pepaya, Cabe Jawa, Kuda laut dan lainnya) sebagai antifertilitas dan fertilitas (afrodisiaka) dengan menggunakan mencit dan tikus sebagai hewan coba. Selain itu, penulis juga pernah meneliti tentang pengaruh faktor genetik dan virus Epstein-Barr pada penderita kanker nasofaring di Indonesia. Pada saat ini penulis sedang melakukan penelitian tentang pengaruh ekstrak kuda laut terhadap ekspresi gen steroidogenesis dan antiinflamasi dengan menggunakan tikus yang diinduksi dengan DMPA dan mencit yang diinduksi dengan lipopolisakarida sebagai hewan model. Di samping itu, juga meneliti tentang studi epigenomik pada pasien sindrom ovarium polikistik (SOPK) di Indonesia yang berfokus pada analisis perbandingan antara pasien SOPK hiperandrogen dan normoandrogen.

Hasil penelitian penulis telah banyak dipublikasikan pada jurnal internasional bereputasi yang terindeks pada *SCImago Journal Rank*, *Scopus*, dan *Web of Science*. Di samping publikasi, penulis juga aktif menjadi *reviewer* pada berbagai jurnal internasional bereputasi, seperti *eCancermedicalscience*, *Food Science & Nutrition*, *International Journal of Cancer*, *Journal of Advanced Veterinary and Animal Research*, *Biodiversity* dan lain sebagainya.



Dr. Andi Muh. Maulana, S.Si., M.Sc., PAK adalah akademisi dan peneliti di bidang ilmu biomedik dengan kepakaran pada anatomi manusia. Beliau menyelesaikan pendidikan doktor pada Program Doktor Ilmu Biomedik FKUI Jakarta, dengan latar riset di bidang onkologi.

Pendidikan magister ditempuh pada Program Studi Ilmu Kedokteran Dasar dan Biomedis FK UGM Yogyakarta, dengan peminatan Anatomi Manusia, sedangkan pendidikan sarjana diselesaikan pada Program Studi Biologi, Universitas Negeri Makassar.

Saat ini, Dr. Andi Muh. Maulana mengabdikan sebagai staf akademik di Fakultas Kedokteran Universitas Muhammadiyah Purwokerto. Beliau menjabat sebagai Ketua Departemen Anatomi serta Ketua Unit Penelitian, Pengabdian kepada Masyarakat, dan Publikasi FK UMP. Dalam peran tersebut, beliau aktif mengoordinasikan pengembangan kegiatan akademik, penelitian, dan publikasi ilmiah di lingkungan fakultas.

Dalam kegiatan pendidikan, beliau mengampu mata kuliah Anatomi serta Metodologi Riset Kedokteran, dengan penekanan pada pemahaman desain studi, pemilihan uji statistik yang tepat, validitas dan reliabilitas data, serta interpretasi hasil penelitian secara kritis dan bertanggung jawab. Pendekatan pengajaran yang dikembangkan menempatkan metodologi dan analisis data sebagai fondasi utama dalam praktik kedokteran berbasis bukti.

Minat keilmuan dan penelitian beliau mencakup bidang onkologi serta eksplorasi potensi bahan alam sebagai agen penghambatan kanker, khususnya kanker payudara. Selain itu, beliau juga berperan sebagai reviewer jurnal internasional bereputasi, terutama untuk naskah yang berkaitan dengan onkologi, biomedik, patologi, serta pengembangan kandidat terapi berbasis bahan alam.

Buku ajar ini disusun sebagai respons terhadap tantangan fundamental yang dihadapi oleh para peneliti, klinisi, dan mahasiswa di bidang kedokteran dasar (Kedokteran) maupun kedokteran klinik, yaitu pemilihan uji statistik yang tepat untuk menganalisis data penelitian. Di era kedokteran berbasis bukti atau evidence-based medicine, kemampuan untuk menerjemahkan data mentah menjadi bukti ilmiah yang valid tidak lagi menjadi keahlian khusus seorang biostatistikawan, melainkan sebuah kompetensi inti bagi setiap insan medis. Kesalahan dalam memilih metode analisis tidak hanya berisiko menghasilkan kesimpulan yang keliru, tetapi juga dapat mendelegitimasi seluruh upaya penelitian, bahkan berpotensi membahayakan keselamatan pasien jika diterapkan dalam praktik klinis. Urgensi penguasaan materi ini semakin meningkat seiring dengan kompleksitas desain riset modern dan volume data yang terus bertambah.

Tujuan utama dari penulisan buku ini adalah untuk menyediakan panduan yang sistematis, logis, dan aplikatif bagi pembaca dalam menavigasi lanskap uji statistik. Buku ini dirancang untuk menjembatani kesenjangan antara pemahaman teoretis tentang statistik dan aplikasinya dalam konteks penelitian Kedokteran yang nyata. Dengan alur yang terstruktur, mulai dari konsep dasar hingga analisis multivariat lanjutan, pembaca akan dibimbing untuk membangun kerangka berpikir analitis. Kerangka ini memungkinkan mereka untuk mengidentifikasi pertanyaan penelitian, mengenali jenis data, memahami asumsi yang mendasari setiap uji, dan pada akhirnya, membuat keputusan metodologis yang dapat dipertanggungjawabkan secara ilmiah. Buku ini diharapkan dapat menjadi sumber daya esensial yang memberdayakan peneliti untuk menghasilkan karya ilmiah yang tidak hanya signifikan secara statistik, tetapi juga relevan secara klinis.