ISBN 978-634-255-715-0



STATISTIK DESKRIPTIF, STATISTIK INFERENSIAL, STATISTIK PARAMETRIK, DAN STATISTIK NONPARAMETRIK

Penulis

Arif Rachman

Timbul Partogi Haposan Simorangkir

STATISTIK DESKRIPTIF, STATISTIK INFERENSIAL, STATISTIK PARAMETRIK, DAN STATISTIK NONPAR

Arif Rachman
Timbul Partogi Haposan Simorangkir

GET PRESS INDONESIA

STATISTIK DESKRIPTIF, STATISTIK INFERENSIAL, STATISTIK PARAMETRIK, DAN STATISTIK NONPAR

Penulis :
Arif Rachman
Timbul Partogi Haposan Simorangkir

ISBN: 978-634-255-715-0

Editor : Dr. Oktavianis, M.Biomed.
Desain Sampul dan Tata Letak : Mila Sari, M. Si.

PENERBIT : GET PRESS INDONESIA
Anggota IKAPI No. 033/SBA/2022
Jl. Palarik RT 01 RW 06 Kelurahan Air Pacah
Kecamatan Koto Tangah Padang Sumatera Barat
website: www.getpress.co.id
email: adm.getpress@gmail.com

Cetakan pertama, Agustus 2026

Hak cipta dilindungi undang-undang.
Dilarang memperbanyak karya tulis ini dalam bentuk
dan dengan cara apapun tanpa izin tertulis dari penerbit.

Undang Undang Republik Indonesia Nomor 19 *Tahun 2002*
Tentang Hak Cipta

Ketentuan Pidana:

Pasal 72

1. Barangsiapa dengan sengaja dan tanpa hak melakukan perbuatan sebagaimana dimaksud dalam Pasal 2 ayat (1) atau Pasal 49 ayat (1) dan ayat (2) dipidana dengan pidana penjara masing-masing paling singkat 1 (satu) bulan dan/atau denda paling sedikit Rp 1.000.000,00 (satu juta rupiah), atau pidana penjara paling lama 7 (tujuh) tahun dan/ atau denda paling banyak Rp 5.000.000.000,00 (lima miliar rupiah).
2. Barangsiapa dengan sengaja menyiarkan, memamerkan, mengedarkan, atau menjual kepada umum suatu Ciptaan atau barang hasil pelanggaran Hak Cipta atau Hak Terkait sebagaimana dimaksud pada ayat (1) dipidana dengan pidana penjara paling lama 5 (lima) tahun dan/ atau denda paling banyak Rp 500.000.000,00 (lima ratus juta rupiah).

KATA PENGANTAR

Puji syukur kepada Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya sehingga buku **"Statistik Deskriptif, Statistik Inferensial, Statistik Parametrik, dan Statistik Nonpar"** ini dapat diselesaikan dengan baik. Buku ini disusun sebagai bahan pembelajaran untuk membantu mahasiswa, peneliti, dan pembaca dalam memahami biostatistik secara sistematis, mulai dari konsep statistik deskriptif dan inferensial, jenis variabel dan skala pengukuran, populasi dan teknik sampling, distribusi data serta uji normalitas, hingga berbagai metode analisis seperti uji perbedaan, korelasi, regresi, analisis data kategorik, diagnostik, dan analisis survival. Selain membahas konsep dasar, buku ini juga memberikan perhatian pada pemilihan metode statistik yang sesuai dengan karakteristik data dan desain penelitian agar hasil analisis dapat diinterpretasikan secara tepat dan bertanggung jawab.

Sejalan dengan perkembangan penelitian dan teknologi analisis data, buku ini juga membahas penggunaan perangkat lunak statistik, *reproducible research*, manajemen data, interpretasi output, pedoman pelaporan ilmiah, serta etika dalam analisis statistik. Materi disusun secara bertahap agar pembaca tidak hanya mampu melakukan analisis statistik, tetapi juga memahami makna hasilnya dalam konteks penelitian dan pengambilan keputusan berbasis bukti. Penulis menyadari bahwa buku ini masih memiliki keterbatasan. Oleh karena itu, kritik dan saran yang konstruktif sangat diharapkan untuk penyempurnaan edisi berikutnya. Semoga buku ini dapat memberikan manfaat bagi mahasiswa, akademisi, peneliti, maupun praktisi dalam meningkatkan kemampuan analisis data dan menghasilkan penelitian yang berkualitas, transparan, serta dapat dipertanggungjawabkan secara ilmiah.

.

Padang, Agustus 2026

Penulis

DAFTAR ISI

KATA PENGANTAR i
DAFTAR ISI ii
DAFTAR TABEL ix
DAFTAR GAMBAR x
BAB 1 KONSEP STATISTIK DESKRIPTIF, STATISTIK INFERENSIAL, STATISTIK PARAMETRIK, DAN STATISTIK NONPARAMETRIK 1
1.1 Pendahuluan 1
1.2 Definisi Statistik Deskriptif 2
1.3 Ukuran Pemusatan Data 3
1.4 Ukuran Penyebaran Data 4
1.5 Visualisasi Data dalam Statistik Deskriptif 5
1.6 Definisi Statistik Inferensial 6
1.7 Statistik Parametrik 8
1.8 Statistik Nonparametrik 9
1.9 Perbandingan Statistik Parametrik dan Nonparametrik 10
1.10 Pemilihan Uji Statistik Berdasarkan Jenis Data dan Desain Penelitian 12
1.11 Asumsi Normalitas dan Central Limit Theorem 13
1.12 Uji Hipotesis, Nilai p, dan Confidence Interval 14
1.13 Kesalahan Umum dalam Penggunaan Statistik Dasar 15
Review Questions 18
DAFTAR PUSTAKA 19
BAB 2 JENIS VARIABEL, SKALA PENGUKURAN, DATA, DAN PERSIAPAN ANALISIS STATISTIK 21
2.1 Pendahuluan 21
2.2 Definisi Variabel dalam Penelitian 22
2.3 Variabel Nominal 23
2.4 Variabel Ordinal 24
2.5 Variabel Interval 25

2.6 Variabel Rasio 27
2.7 Variabel Kategorik dan Variabel Numerik 28
2.8 Variabel Diskrit dan Kontinu 29
2.9 Variabel Independen dan Dependen 30
2.10 Kovariat, Confounder, Moderator, dan Mediator 31
2.11 Coding Data dan Struktur Dataset 32
2.12 Data Missing, Outlier, dan Kesalahan Entri 33
2.13 Transformasi Data dan Pengelompokan Variabel 34
2.14 Hubungan Jenis Variabel dengan Statistik Deskriptif 36
2.15 Hubungan Jenis Variabel dengan Pemilihan Uji Statistik 37
Review Questions 40
DAFTAR PUSTAKA 41
BAB 3 POPULASI, SAMPEL, TEKNIK SAMPLING, UKURAN SAMPEL, DAN KEKUATAN STATISTIK 43
3.1 Pendahuluan 43
3.2 Populasi Target dan Populasi Terjangkau 44
3.3 Sampling Frame dan Unit Analisis 45
3.4 Sampel dan Representativitas 47
3.5 Teknik Sampling Probabilitas 48
3.6 Simple Random Sampling dan Systematic Sampling 49
3.7 Stratified Sampling dan Cluster Sampling 50
3.8 Teknik Sampling Nonprobabilitas 52
3.9 Bias Sampling dan Validitas Eksternal 53
3.10 Ukuran Sampel dan Presisi Estimasi 54
3.11 Power, Kesalahan Tipe I, dan Kesalahan Tipe II 56
3.12 Effect Size dan Variabilitas Data 57
3.13 Ukuran Sampel untuk Studi Deskriptif dan Estimasi Prevalensi 58
3.14 Ukuran Sampel untuk Studi Komparatif 59
3.15 Ukuran Sampel untuk Studi Korelasi dan Regresi 61

3.16 Dropout, Missing Data, dan Inflasi Ukuran Sampel 62
3.17 Kesalahan Umum dalam Sampling dan Ukuran Sampel 63
Review Questions 66
DAFTAR PUSTAKA 68
BAB 4 DISTRIBUSI DATA, UJI NORMALITAS, TRANSFORMASI DATA, DAN PEMILIHAN STATISTIK PARAMETRIK–NONPARAMETRIK 69
4.1 Pendahuluan 69
4.2 Konsep Distribusi Data 70
4.3 Distribusi Normal 71
4.4 Distribusi Tidak Normal dan Skewness 72
4.5 Kurtosis dan Bentuk Puncak Distribusi 73
4.6 Outlier dan Pengaruhnya terhadap Analisis 74
4.7 Histogram sebagai Pemeriksaan Distribusi 76
4.8 Q–Q Plot dan Pemeriksaan Visual Normalitas 77
4.9 Uji Shapiro–Wilk dan Kolmogorov–Smirnov 78
4.10 Central Limit Theorem 79
4.11 Transformasi Data 80
4.12 Statistik Robust dan Alternatif Analisis 81
4.13 Pemilihan Uji Parametrik dan Nonparametrik Berdasarkan Distribusi 83
4.14 Normalitas pada Regresi dan Residual 84
4.15 Kesalahan Umum dalam Menilai Normalitas 85
4.16 Pelaporan Distribusi Data dalam Artikel Ilmiah 86
Review Questions 89
DAFTAR PUSTAKA 90
BAB 5 UJI PERBEDAAN KELOMPOK: UJI T, MANN–WHITNEY, WILCOXON, ANOVA, DAN KRUSKAL–WALLIS 91
5.1 Pendahuluan 91
5.2 Konsep Dasar Uji Perbedaan Kelompok 92
5.3 Independent t-test 94

5.4 Mann–Whitney U Test 96
5.5 Paired t-test 98
5.6 Wilcoxon Signed-Rank Test 100
5.7 One-Way ANOVA 102
5.8 Kruskal–Wallis Test 104
5.9 Repeated Measures ANOVA dan Friedman Test 106
5.10 Uji Chi-Square dan Fisher's Exact untuk Data Kategorik 108
5.11 Post-Hoc Test dan Multiple Comparisons 110
5.12 Effect Size dalam Uji Perbedaan Kelompok 112
5.13 Pelaporan Hasil Uji Perbedaan Kelompok 114
5.14 Kesalahan Umum dalam Uji Perbedaan Kelompok 116
Review Questions 120
DAFTAR PUSTAKA 121
BAB 6 ANALISIS KORELASI DAN REGRESI: PEARSON, SPEARMAN, REGRESI LINEAR, REGRESI LOGISTIK, DAN INTERPRETASI MODEL STATISTIK 123
6.1 Pendahuluan 123
6.2 Konsep Dasar Asosiasi Antarvariabel 124
6.3 Scatterplot sebagai Langkah Awal Analisis Hubungan 125
6.4 Korelasi Pearson 126
6.5 Korelasi Spearman 127
6.6 Korelasi Kendall 129
6.7 Interpretasi Koefisien Korelasi 130
6.8 Korelasi Tidak Sama dengan Kausalitas 130
6.9 Regresi Linear Sederhana 131
6.10 Regresi Linear Berganda 133
6.11 Asumsi Regresi Linear 134
6.12 Multikolinearitas dalam Regresi 135
6.13 Regresi Logistik 136
6.14 Odds Ratio dalam Regresi Logistik 137

6.15 Evaluasi Model Regresi Logistik 138
6.16 Overfitting dan Validasi Model 139
6.17 Confounding dan Adjustment dalam Regresi 140
6.18 Pelaporan Hasil Korelasi dan Regresi 141
6.19 Kesalahan Umum dalam Analisis Korelasi dan Regresi 142
Review Questions 144
DAFTAR PUSTAKA 145
BAB 7 ANALISIS DATA KATEGORIK DAN DIAGNOSTIK: TABEL KONTINGENSI, CHI-SQUARE, FISHER'S EXACT, MCNEMAR, SENSITIVITAS, SPESIFISITAS, DAN ROC CURVE 147
7.1 Pendahuluan 147
7.2 Jenis Data Kategorik 148
7.3 Tabel Kontingensi 149
7.4 Chi-Square Test 150
7.5 Fisher's Exact Test 151
7.6 McNemar Test untuk Data Berpasangan 152
7.7 Risk Difference, Risk Ratio, dan Odds Ratio 154
7.8 Analisis Diagnostik dan Tabel 2×2 154
7.9 Sensitivitas dan Spesifisitas 155
7.10 Nilai Prediksi Positif dan Nilai Prediksi Negatif 156
7.11 Likelihood Ratio dan Accuracy 157
7.12 ROC Curve dan AUC 158
7.13 Pemilihan Cut-Off Diagnostik 159
7.14 Analisis Diagnostik pada Model Prediksi Klinis 161
7.15 Pelaporan Analisis Data Kategorik 162
7.16 Kesalahan Umum dalam Analisis Kategorik dan Diagnostik 163
Review Questions 166
DAFTAR PUSTAKA 168
BAB 8 ANALISIS SURVIVAL, MODEL PREDIKSI KLINIS, DAN PEDOMAN PELAPORAN STATISTIK 169

8.1 Pendahuluan 169
8.2 Konsep Dasar Data Time-to-Event 170
8.3 Censoring dalam Analisis Survival 171
8.4 Kaplan–Meier Curve 172
8.5 Log-Rank Test 174
8.6 Cox Proportional Hazards Regression 175
8.7 Hazard Ratio dan Interpretasinya 176
8.8 Asumsi Proportional Hazards 177
8.9 Model Prediksi Klinis 179
8.10 Diskriminasi, Kalibrasi, dan Validasi Model 180
8.11 CONSORT untuk Randomized Trials 181
8.12 STARD untuk Studi Akurasi Diagnostik 182
8.13 PRISMA untuk Systematic Review dan Meta-Analysis 184
8.14 Pelaporan Nilai p, Confidence Interval, dan Effect Size 185
8.15 Kesalahan Umum dalam Analisis Survival dan Pelaporan Statistik 186
Review Questions 190
DAFTAR PUSTAKA 191
BAB 9 APLIKASI PERANGKAT LUNAK STATISTIK, REPRODUCIBLE RESEARCH, MANAJEMEN DATA, DAN INTERPRETASI OUTPUT ANALISIS 193
9.1 Pendahuluan 193
9.2 Peran Perangkat Lunak Statistik dalam Penelitian 194
9.3 Struktur Dataset dan Codebook 196
9.4 Input Data dan Validasi Awal 197
9.5 Cleaning Data dan Penanganan Missing Value 198
9.6 Coding, Recode, dan Transformasi Variabel 199
9.7 Analisis Deskriptif dengan Perangkat Lunak Statistik 200
9.8 Analisis Inferensial dengan Software 202
9.9 Sintaks, Script, dan Reproducible Workflow 203

9.10 Interpretasi Output SPSS, R, jamovi, dan JASP 204
9.11 Visualisasi Data dan Kualitas Grafik Ilmiah 205
9.12 Audit Analisis dan Quality Control 207
9.13 Pelaporan Output Statistik dalam Artikel Ilmiah 208
9.14 Kesalahan Umum dalam Penggunaan Software Statistik 209
Review Questions Bab 9 213
DAFTAR PUSTAKA 215
BAB 10 SINTESIS AKHIR BIOSTATISTIK: INTEGRASI DESAIN PENELITIAN, ANALISIS DATA, INTERPRETASI STATISTIK, PELAPORAN ILMIAH, DAN ETIKA REPRODUCIBLE RESEARCH 217
10.1 Pendahuluan 217
10.2 Dari Pertanyaan Penelitian ke Rencana Analisis Statistik 218
10.3 Integrasi Statistik Deskriptif dan Inferensial 219
10.4 Pemilihan Uji Statistik Berdasarkan Data dan Desain 220
10.5 Nilai p, Confidence Interval, dan Effect Size 222
10.6 Bias, Confounding, dan Validitas Penelitian 223
10.7 Reproducible Research dan FAIR Data 224
10.8 Pedoman Pelaporan Berdasarkan Desain Studi 225
10.9 Statistik dalam Era Artificial Intelligence dan Model Prediksi 226
10.10 Etika Statistik dan Tanggung Jawab Peneliti 227
10.11 Audit Statistik sebelum Publikasi 228
10.12 Model Praktis Akhir: Sepuluh Langkah Analisis Statistik 229
Review Questions 233
DAFTAR PUSTAKA 234

DAFTAR TABEL

Tabel 1. 1. Perbandingan statistik parametrik dan nonparametrik 11
Tabel 2. 1. Perbandingan variabel kategorik dan numerik 29
Tabel 4. 1. Pemilihan uji berdasarkan distribusi dan desain penelitian 83
Tabel 6. 1. Interpretasi umum korelasi 130
Tabel 7. 1. Contoh tabel kontingensi 2×2 150
Tabel 7. 2. Ukuran asosiasi pada data kategorik 154
Tabel 9. 1. Elemen codebook penelitian 196
Tabel 9. 2. Perbandingan umum perangkat lunak statistik 205
Tabel 10. 1. Pedoman pelaporan utama dalam penelitian kesehatan 226

DAFTAR GAMBAR

Gambar 1. 1. Peta konsep statistik dasar dalam penelitian 2
Gambar 1. 2. Komponen statistik deskriptif 3
Gambar 1. 3. Mean, median, dan modus pada distribusi data 4
Gambar 1. 4. Ukuran penyebaran data 5
Gambar 1. 5. Visualisasi data statistik deskriptif 6
Gambar 1. 6. Alur statistik inferensial 7
Gambar 1. 7. Statistik parametrik dan asumsi utamanya 9
Gambar 1. 8. Statistik nonparametrik dan kondisi penggunaannya 10
Gambar 1. 9. Pohon keputusan sederhana pemilihan uji statistic 13
Gambar 1. 10. Distribusi normal, distribusi tidak normal, dan Q–Q plot 14
Gambar 1. 11. Uji hipotesis dan confidence interval 15
Gambar 1. 12. Kesalahan umum dalam statistik dasar 16
Gambar 1. 13. Ringkasan konsep statistik dasar 17
Gambar 2. 1. Peta konsep jenis variabel dan skala pengukuran 22
Gambar 2. 2. Definisi variabel dan operasionalisasi variabel 23
Gambar 2. 3. Variabel nominal dalam penelitian Kesehatan 24
Gambar 2. 4. Variabel ordinal dan analisis nonparametric 25
Gambar 2. 5. Variabel interval dan interpretasi statistic 26
Gambar 2. 6. Variabel rasio dalam data klinis 28
Gambar 2. 7. Variabel diskrit dan kontinu 30
Gambar 2. 8. Variabel independen dan dependen dalam penelitian 31
Gambar 2. 9. Kovariat, confounder, moderator, dan mediator 32
Gambar 2. 10. Struktur dataset dan codebook penelitian 33
Gambar 2. 11. Data missing, outlier, dan kesalahan entri 34
Gambar 2. 12. Transformasi dan kategorisasi variable 35
Gambar 2. 13. Pemilihan uji statistik berdasarkan variabel outcome 38
Gambar 2. 14. Ringkasan jenis variabel dan persiapan analisis statistic 39
Gambar 3. 1. Hubungan populasi, sampel, sampling, dan inferensi statistic 44
Gambar 3. 2. Populasi target dan populasi terjangkau 45
Gambar 3. 3. Sampling frame dan unit analisis 46
Gambar 3. 4. Representativitas sampel dalam penelitian 48
Gambar 3. 5. Jenis teknik sampling probabilitas 49
Gambar 3. 6. Simple random sampling dan systematic sampling 50

Gambar 3. 7. Stratified sampling dan cluster sampling 51
Gambar 3. 8. Teknik sampling nonprobabilitas 53
Gambar 3. 9. Bias sampling dan validitas eksternal 54
Gambar 3. 10. Faktor yang memengaruhi ukuran sampel 55
Gambar 3. 11. Power dan kesalahan statistic 56
Gambar 3. 12. Effect size, variasi data, dan ukuran sampel 58
Gambar 3. 13. Ukuran sampel untuk studi prevalensi dan estimasi rata-rata 59
Gambar 3. 14. Ukuran sampel pada studi komparatif 60
Gambar 3. 15. Ukuran sampel untuk korelasi dan regresi 62
Gambar 3. 16. Dropout dan inflasi ukuran sampel 63
Gambar 3. 17. Kesalahan umum dalam sampling dan ukuran sampel 64
Gambar 3. 18. Ringkasan populasi, sampel, sampling, dan ukuran sampel 65
Gambar 4. 1. Peta konsep distribusi data dan pemilihan uji statistic 70
Gambar 4. 2. Bentuk umum distribusi data 71
Gambar 4. 3. Distribusi normal dan posisi mean–median–modus 72
Gambar 4. 4. Skewness positif dan negative 73
Gambar 4. 5. Kurtosis dalam distribusi data 74
Gambar 4. 6. Outlier dan strategi penanganannya 75
Gambar 4. 7. Histogram untuk menilai distribusi data 76
Gambar 4. 8. Transformasi data menceng 81
Gambar 4. 9. Alternatif saat asumsi parametrik tidak terpenuhi 82
Gambar 4. 10. Pemeriksaan residual dalam regresi linear 85
Gambar 4. 11. Pelaporan distribusi data dan pilihan analisis 87
Gambar 4. 12. Ringkasan distribusi data dan keputusan analisis statistic 88
Gambar 5. 1. Peta konsep uji perbedaan kelompok 92
Gambar 5. 2. Dasar pemilihan uji perbedaan kelompok 93
Gambar 5. 3. Independent t-test untuk dua kelompok independent 95
Gambar 5. 4. Mann–Whitney U test berbasis peringkat 97
Gambar 5. 5. Paired t-test pada desain sebelum–sesudah 99
Gambar 5. 6. Wilcoxon signed-rank test 101
Gambar 5. 7. One-way ANOVA untuk lebih dari dua kelompok 103
Gambar 5. 8. Kruskal–Wallis test untuk lebih dari dua kelompok 105
Gambar 5. 9. Uji lebih dari dua pengukuran berulang 107
Gambar 5. 10. Uji untuk data kategorik 109
Gambar 5. 11. Post-hoc test dan multiple comparisons 111
Gambar 5. 12. Effect size dan nilai p 113

Gambar 5. 13. Format pelaporan uji perbedaan kelompok 115
Gambar 5. 14. Kesalahan umum dalam uji perbedaan kelompok 117
Gambar 5. 15. Ringkasan uji perbedaan kelompok 119
Gambar 6. 1. Peta konsep korelasi dan regresi dalam penelitian Kesehatan 124
Gambar 6. 2. Asosiasi, korelasi, dan kausalitas 125
Gambar 6. 3. Scatterplot dalam analisis korelasi dan regresi 126
Gambar 6. 4. Korelasi Pearson 127
Gambar 6. 5. Korelasi Spearman berbasis peringkat 128
Gambar 6. 6. Korelasi Kendall dan pasangan concordant–discordant 129
Gambar 6. 7. Korelasi, confounding, dan kausalitas 131
Gambar 6. 8. Regresi linear sederhana 132
Gambar 6. 9. Regresi linear berganda dalam penelitian klinis 133
Gambar 6. 10. Pemeriksaan asumsi regresi linear 134
Gambar 6. 11. Multikolinearitas dalam regresi berganda 135
Gambar 6. 12. Regresi logistik untuk outcome biner 136
Gambar 6. 13. Odds ratio dan interpretasi regresi logistic 137
Gambar 6. 14. Evaluasi model regresi logistik 138
Gambar 6. 15. Overfitting dan validasi model prediksi 139
Gambar 6. 16. Confounding dan adjustment dalam regresi 140
Gambar 6. 17. Pelaporan korelasi dan regresi 141
Gambar 6. 18. Kesalahan umum korelasi dan regresi 142
Gambar 6. 19. Ringkasan analisis korelasi dan regresi 143
Gambar 7. 1. Peta konsep analisis data kategorik dan diagnostic 148
Gambar 7. 2. Klasifikasi data kategorik 149
Gambar 7. 3. Prinsip chi-square test 151
Gambar 7. 4. Fisher's exact test pada tabel 2×2 152
Gambar 7. 5. McNemar test pada data sebelum–sesudah 153
Gambar 7. 6. Sensitivitas dan spesifisitas 156
Gambar 7. 7. Nilai prediksi dan pengaruh prevalensi 157
Gambar 7. 8. Likelihood ratio dan accuracy 158
Gambar 7. 9. ROC curve dan AUC 159
Gambar 7. 10. Pemilihan cut-off diagnostic 160
Gambar 7. 11. Evaluasi diagnostik model prediksi klinis 162
Gambar 7. 12. Pelaporan analisis kategorik dan diagnostic 163
Gambar 7. 13. Kesalahan umum analisis kategorik dan diagnostic 164
Gambar 7. 14. Ringkasan analisis data kategorik dan diagnostic 165
Gambar 8. 1. Peta konsep analisis survival dan pelaporan statistic 170

Gambar 8. 2. Struktur data time-to-event 171
Gambar 8. 3. Censoring pada data survival 172
Gambar 8. 4. Kaplan–Meier curve 173
Gambar 8. 5. Log-rank test untuk membandingkan survival curve 175
Gambar 8. 6. Cox proportional hazards regression 176
Gambar 8. 7. Interpretasi hazard ratio 177
Gambar 8. 8. Pemeriksaan proportional hazards assumption 178
Gambar 8. 9. Model prediksi klinis 180
Gambar 8. 10. Diskriminasi, kalibrasi, dan validasi model prediksi 181
Gambar 8. 11. Diagram alur CONSORT dalam randomized trial 182
Gambar 8. 12. STARD dalam studi akurasi diagnostic 183
Gambar 8. 13. PRISMA, forest plot, dan funnel plot 185
Gambar 8. 14. Pelaporan hasil statistik yang lengkap 186
Gambar 8. 15. Kesalahan umum analisis survival dan pelaporan statistic 187
Gambar 8. 16. Ringkasan analisis survival dan pedoman pelaporan 189
Gambar 9. 1. Peta konsep perangkat lunak statistik dan reproducible research 194
Gambar 9. 2. Peran perangkat lunak statistik dalam workflow penelitian 195
Gambar 9. 3. Validasi awal dataset 198
Gambar 9. 4. Cleaning data dan missing value 199
Gambar 9. 5. Coding, recode, dan transformasi variable 200
Gambar 9. 6. Analisis deskriptif berbasis software 201
Gambar 9. 7. Interpretasi output inferensial 203
Gambar 9. 8. Reproducible workflow dalam analisis statistic 204
Gambar 9. 9. Visualisasi data ilmiah berbasis software 206
Gambar 9. 10. Audit dan quality control analisis statistic 208
Gambar 9. 11. Pelaporan output statistik dalam artikel ilmiah 209
Gambar 9. 12. Kesalahan umum penggunaan software statistic 210
Gambar 9. 13. Ringkasan perangkat lunak statistik dan reproducible research 212
Gambar 10. 1. Sintesis akhir proses biostatistik dalam penelitian Kesehatan 218
Gambar 10. 2. Alur dari pertanyaan penelitian ke rencana analisis statistic 219
Gambar 10. 3. Integrasi statistik deskriptif dan inferensial 220
Gambar 10. 4. Matriks pemilihan uji statistik berdasarkan desain dan jenis data 221

Gambar 10. 5. Interpretasi nilai p, confidence interval, dan effect size 222
Gambar 10. 6. Bias, confounding, dan validitas penelitian 224
Gambar 10. 7. Reproducible workflow dan FAIR data dalam
penelitian Kesehatan 225
Gambar 10. 8. Statistik, AI, dan model prediksi klinis 227
Gambar 10. 9. Etika statistik dalam penelitian biomedis 228
Gambar 10. 10. Audit statistik sebelum publikasi 229
Gambar 10. 11. Model sepuluh langkah analisis statistik dalam
penelitian Kesehatan 230
Gambar 10. 12. Ringkasan akhir biostatistik terapan 232

BAB 1
KONSEP STATISTIK DESKRIPTIF, STATISTIK INFERENSIAL, STATISTIK PARAMETRIK, DAN STATISTIK NONPARAMETRIK

1.1 Pendahuluan

Statistik merupakan perangkat ilmiah yang digunakan untuk mengubah data mentah menjadi informasi yang bermakna, terstruktur, dan dapat digunakan dalam pengambilan keputusan penelitian (Cooksey, 2020). Statistik membantu peneliti memahami pola data, mengukur variasi, membandingkan kelompok, menganalisis hubungan antarvariabel, dan menafsirkan ketidakpastian hasil penelitian (Kaliyadan & Kulkarni, 2019). Statistik secara umum dapat dibagi menjadi statistik deskriptif dan statistik inferensial (Cooksey, 2020). Statistik deskriptif berfokus pada peringkasan, pengorganisasian, dan penyajian data yang diamati dalam suatu sampel atau dataset (Cooksey, 2020). Statistik inferensial digunakan untuk menarik kesimpulan dari sampel ke populasi melalui estimasi, uji hipotesis, confidence interval, dan nilai p (Rovetta, 2024).

Statistik inferensial dapat dilakukan dengan pendekatan parametrik atau nonparametrik sesuai dengan jenis data, distribusi data, asumsi analisis, dan desain penelitian (Chicco *et al*., 2025). Statistik parametrik umumnya digunakan pada data numerik berskala interval atau rasio ketika asumsi tertentu seperti normalitas dan homogenitas varians terpenuhi (Ghasemi & Zahediasl, 2012). Statistik nonparametrik digunakan ketika data tidak memenuhi asumsi parametrik, berskala ordinal atau nominal, berukuran sampel kecil, menceng, atau mengandung outlier ekstrem (Chicco *et al*., 2025). Pemilihan metode statistik harus disesuaikan dengan pertanyaan penelitian, jenis variabel, jumlah kelompok, hubungan antarobservasi, distribusi data, dan tujuan analisis (Kaliyadan & Kulkarni, 2019).



Gambar 1. 1. Peta konsep statistik dasar dalam penelitian

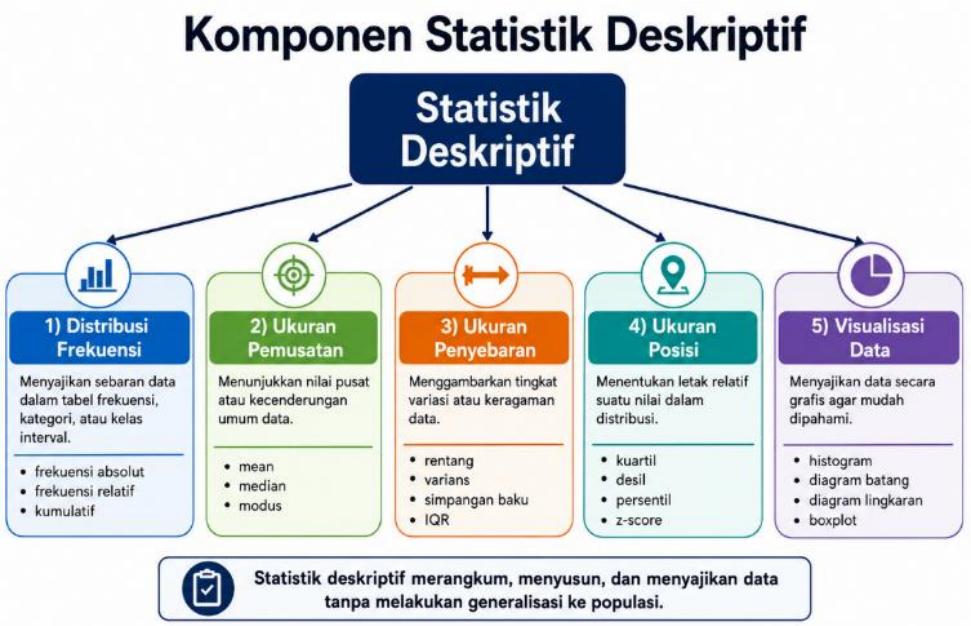
Ilustrasi menunjukkan hubungan antara data, statistik deskriptif, statistik inferensial, statistik parametrik, dan statistik nonparametrik dalam proses penelitian kuantitatif. **Sumber:** Diadaptasi dari Cooksey (2020), Kaliyadan dan Kulkarni (2019), Chicco *et al*. (2025), serta Ghasemi dan Zahediasl (2012).

1.2 Definisi Statistik Deskriptif

Statistik deskriptif adalah cabang statistik yang digunakan untuk mengorganisasi, meringkas, menggambarkan, dan menyajikan data tanpa melakukan generalisasi langsung ke populasi yang lebih luas (Cooksey, 2020). Statistik deskriptif digunakan untuk menjawab pertanyaan tentang karakteristik dasar data, seperti nilai rata-rata, nilai tengah, frekuensi kategori, variasi data, bentuk distribusi, dan keberadaan outlier (Kaliyadan & Kulkarni, 2019). Data kategorik biasanya diringkas dengan frekuensi dan persentase (Kaliyadan & Kulkarni, 2019). Data numerik biasanya diringkas dengan ukuran pemusatan, ukuran penyebaran, ukuran posisi, dan grafik distribusi (Cooksey, 2020).

Statistik deskriptif tidak digunakan untuk membuktikan hipotesis atau menarik kesimpulan kausal (Cooksey, 2020). Statistik deskriptif tetap penting karena memberikan gambaran awal mengenai karakteristik sampel sebelum peneliti melakukan analisis inferensial (Kaliyadan &

Kulkarni, 2019). Statistik deskriptif juga membantu peneliti memeriksa kualitas data, mendeteksi nilai ekstrem, memahami pola distribusi, dan memilih uji statistik yang sesuai (Ghasemi & Zahediasl, 2012). Penyajian statistik deskriptif yang baik harus menyesuaikan ukuran ringkasan dengan skala data dan bentuk distribusi data (Kaliyadan & Kulkarni, 2019).



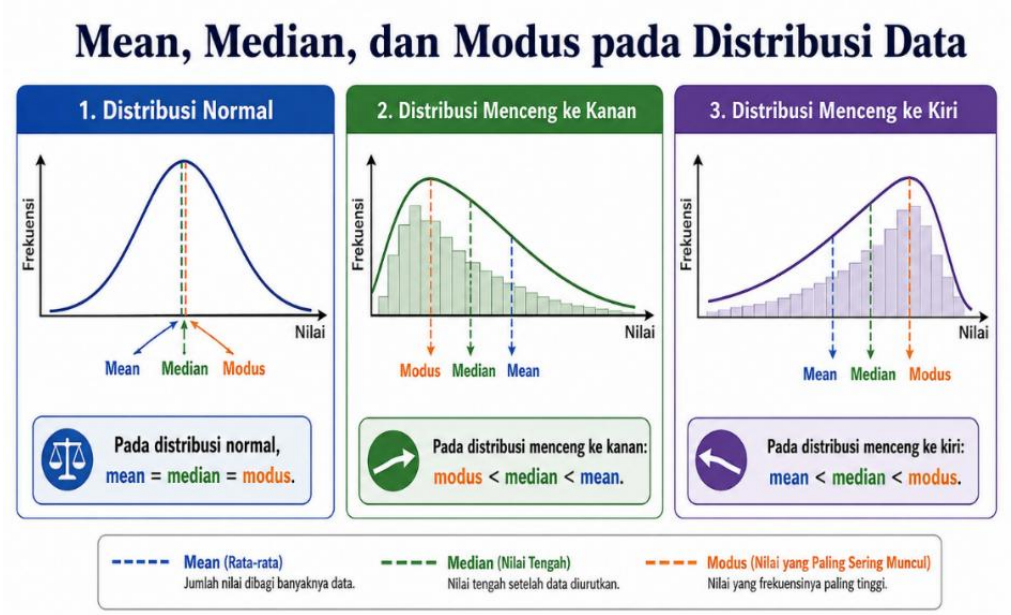
Gambar 1. 2. Komponen statistik deskriptif

Diagram menunjukkan statistik deskriptif sebagai gabungan distribusi frekuensi, ukuran pemusatan, ukuran penyebaran, ukuran posisi, dan visualisasi data. **Sumber:** Diadaptasi dari Cooksey (2020) dan Kaliyadan dan Kulkarni (2019).

1.3 Ukuran Pemusatan Data

Ukuran pemusatan data menunjukkan nilai yang mewakili pusat atau lokasi utama suatu distribusi data (Cooksey, 2020). Ukuran pemusatan yang paling sering digunakan adalah mean, median, dan modus (Kaliyadan & Kulkarni, 2019). Mean atau rata-rata dihitung dengan menjumlahkan seluruh nilai observasi lalu membaginya dengan jumlah observasi (Cooksey, 2020). Median adalah nilai tengah setelah data diurutkan dari nilai terkecil ke nilai terbesar (Kaliyadan & Kulkarni, 2019). Modus adalah nilai atau kategori yang paling sering muncul dalam dataset (Kaliyadan & Kulkarni, 2019).

Pemilihan ukuran pemusatan harus disesuaikan dengan bentuk distribusi data (Cooksey, 2020). Mean lebih tepat digunakan ketika data numerik berdistribusi relatif simetris dan tidak memiliki outlier ekstrem (Ghasemi & Zahediasl, 2012). Median lebih tepat digunakan ketika data menceng atau memiliki nilai ekstrem karena median lebih tahan terhadap pengaruh outlier (Cooksey, 2020). Modus berguna untuk data kategorik karena dapat menunjukkan kategori yang paling dominan dalam dataset (Kaliyadan & Kulkarni, 2019). Penggunaan mean tanpa memeriksa distribusi data dapat menyesatkan apabila data sangat menceng atau memiliki outlier ekstrem (Ghasemi & Zahediasl, 2012).



Gambar 1. 3. Mean, median, dan modus pada distribusi data

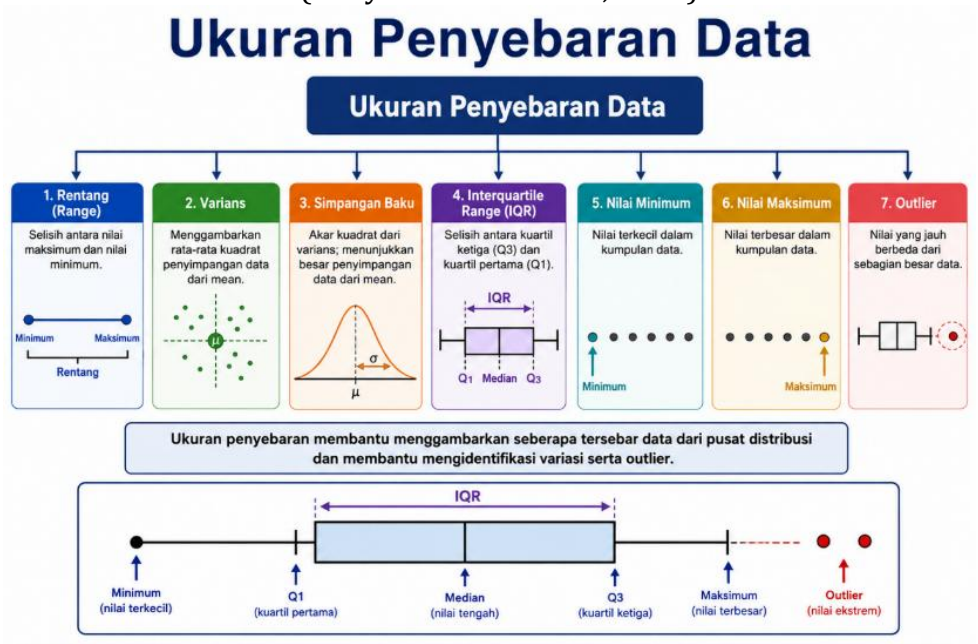
Gambar menunjukkan perbedaan posisi mean, median, dan modus pada distribusi normal, distribusi menceng ke kanan, dan distribusi menceng ke kiri. **Sumber:** Diadaptasi dari Cooksey (2020), Kaliyadan dan Kulkarni (2019), serta Ghasemi dan Zahediasl (2012).

1.4 Ukuran Penyebaran Data

Ukuran penyebaran data menjelaskan seberapa jauh nilai-nilai observasi tersebar dari pusat distribusi (Cooksey, 2020). Ukuran penyebaran yang sering digunakan adalah rentang, varians, simpangan baku, interquartile range, nilai minimum, dan nilai maksimum (Kaliyadan & Kulkarni, 2019). Simpangan baku sering dilaporkan bersama mean ketika data berdistribusi normal atau mendekati normal (Ghasemi &

Zahediasl, 2012). Interquartile range sering dilaporkan bersama median ketika data menceng atau mengandung outlier (Cooksey, 2020).

Penyebaran data penting karena dua dataset dapat memiliki mean yang sama tetapi variasi data yang sangat berbeda (Cooksey, 2020). Dataset dengan variasi kecil menunjukkan bahwa sebagian besar nilai berada dekat pusat distribusi (Cooksey, 2020). Dataset dengan variasi besar menunjukkan bahwa nilai observasi tersebar lebih luas dan mungkin memiliki heterogenitas yang lebih tinggi (Kaliyadan & Kulkarni, 2019). Laporan statistik deskriptif yang hanya menyajikan ukuran pemusatan tanpa ukuran penyebaran menjadi kurang informatif (Cooksey, 2020). Statistik deskriptif yang baik sebaiknya menyajikan ukuran pemusatan dan ukuran penyebaran secara berpasangan sesuai bentuk distribusi data (Kaliyadan & Kulkarni, 2019).



Gambar 1. 4. Ukuran penyebaran data

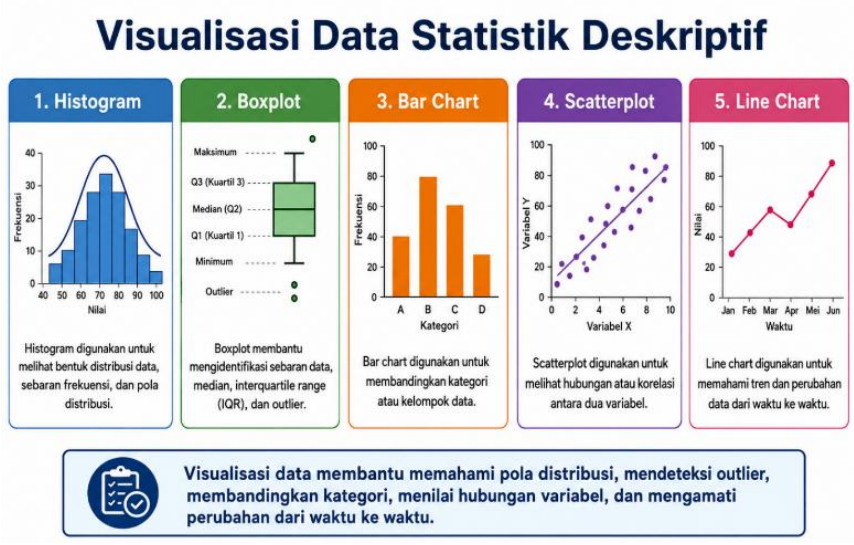
Diagram menunjukkan rentang, varians, simpangan baku, interquartile range, nilai minimum, nilai maksimum, dan outlier dalam distribusi data. **Sumber:** Diadaptasi dari Cooksey (2020) dan Kaliyadan dan Kulkarni (2019).

1.5 Visualisasi Data dalam Statistik Deskriptif

Visualisasi data merupakan bagian penting dari statistik deskriptif karena grafik dapat menunjukkan pola data yang tidak selalu tampak dari angka ringkasan (Cooksey, 2020). Histogram digunakan untuk menilai bentuk distribusi data numerik (Ghasemi & Zahediasl, 2012). Boxplot

digunakan untuk melihat median, kuartil, sebaran, dan outlier (Cooksey, 2020). Bar chart digunakan untuk menyajikan data kategorik dalam bentuk frekuensi atau persentase (Kaliyadan & Kulkarni, 2019). Scatterplot digunakan untuk menggambarkan hubungan antara dua variabel numerik (Cooksey, 2020).

Visualisasi data juga membantu peneliti memeriksa asumsi statistik sebelum melakukan analisis inferensial (Ghasemi & Zahediasl, 2012). Histogram dan Q–Q plot dapat membantu menilai apakah data mendekati distribusi normal (Ghasemi & Zahediasl, 2012). Boxplot dapat membantu mengidentifikasi outlier yang berpotensi memengaruhi mean dan uji parametrik (Cooksey, 2020). Scatterplot dapat membantu menilai apakah hubungan dua variabel bersifat linear atau tidak linear (Cooksey, 2020). Oleh karena itu, visualisasi data harus dipandang sebagai bagian dari validasi awal sebelum analisis statistik lanjutan (Ghasemi & Zahediasl, 2012).



Gambar 1. 5. Visualisasi data statistik deskriptif

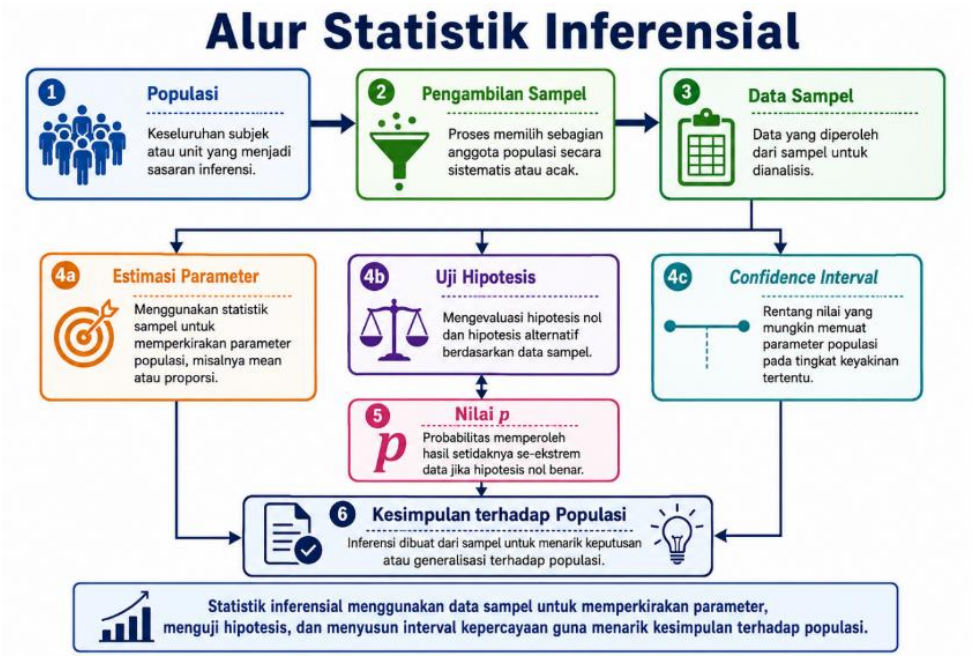
Gambar menunjukkan histogram, boxplot, bar chart, scatterplot, dan line chart sebagai alat untuk memahami pola distribusi, outlier, kategori, hubungan variabel, dan perubahan waktu. **Sumber:** Diadaptasi dari Cooksey (2020), Kaliyadan dan Kulkarni (2019), serta Ghasemi dan Zahediasl (2012).

1.6 Definisi Statistik Inferensial

Statistik inferensial adalah cabang statistik yang digunakan untuk menarik kesimpulan tentang populasi berdasarkan data sampel (Rovetta, 2024). Statistik inferensial digunakan untuk memperkirakan parameter

populasi, menguji hipotesis, menghitung confidence interval, menilai nilai p, dan mengevaluasi ketidakpastian hasil penelitian (Rovetta, 2024). Statistik inferensial bekerja berdasarkan teori probabilitas dan distribusi sampling (Rovetta, 2024). Penggunaan statistik inferensial harus selalu dikaitkan dengan desain penelitian, kualitas sampel, dan konteks ilmiah penelitian (Rovetta, 2024).

Perbedaan utama statistik deskriptif dan statistik inferensial terletak pada ruang lingkup kesimpulan (Cooksey, 2020). Statistik deskriptif hanya menggambarkan data yang diamati dalam dataset (Cooksey, 2020). Statistik inferensial mencoba menarik kesimpulan dari data sampel ke populasi yang lebih luas (Rovetta, 2024). Statistik inferensial selalu mengandung ketidakpastian karena sampel tidak pernah sepenuhnya identik dengan populasi (Rovetta, 2024). Oleh karena itu, hasil inferensial harus ditafsirkan melalui estimasi efek, confidence interval, nilai p, dan konteks ilmiah (Rovetta, 2024).



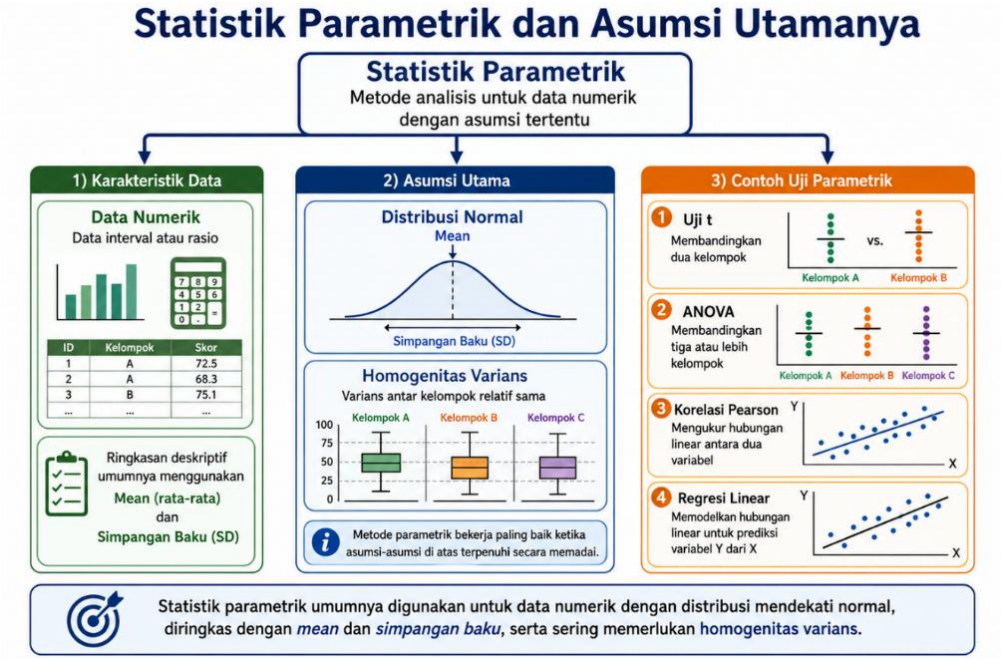
Gambar 1. 6. Alur statistik inferensial

Diagram menunjukkan populasi, pengambilan sampel, data sampel, estimasi parameter, uji hipotesis, confidence interval, nilai p, dan kesimpulan terhadap populasi. **Sumber:** Diadaptasi dari Rovetta (2024), Chicco *et al*. (2025), serta Kaliyadan dan Kulkarni (2019).

1.7 Statistik Parametrik

Statistik parametrik adalah metode statistik yang menggunakan asumsi tertentu tentang distribusi data atau parameter populasi (Ghasemi & Zahediasl, 2012). Statistik parametrik umumnya digunakan untuk data numerik berskala interval atau rasio (Kaliyadan & Kulkarni, 2019). Statistik parametrik sering mensyaratkan asumsi normalitas, homogenitas varians, independensi observasi, dan hubungan linear pada analisis tertentu (Ghasemi & Zahediasl, 2012). Contoh uji parametrik adalah independent t-test, paired t-test, one-way ANOVA, repeated measures ANOVA, Pearson correlation, dan linear regression (Chicco *et al*., 2025).

Uji parametrik dapat memiliki kekuatan statistik lebih tinggi dibanding uji nonparametrik ketika asumsi analisis terpenuhi (Chicco *et al*., 2025). Normalitas data perlu dievaluasi sebelum menggunakan banyak uji parametrik, terutama pada ukuran sampel kecil (Ghasemi & Zahediasl, 2012). Central limit theorem menjelaskan bahwa distribusi sampling dari mean dapat mendekati normal ketika ukuran sampel cukup besar meskipun data mentah tidak sepenuhnya normal (Ghasemi & Zahediasl, 2012). Namun, central limit theorem tidak menghilangkan kebutuhan untuk memeriksa outlier, kemencengan ekstrem, dan kesesuaian model statistik (Ghasemi & Zahediasl, 2012).



Gambar 1. 7. Statistik parametrik dan asumsi utamanya

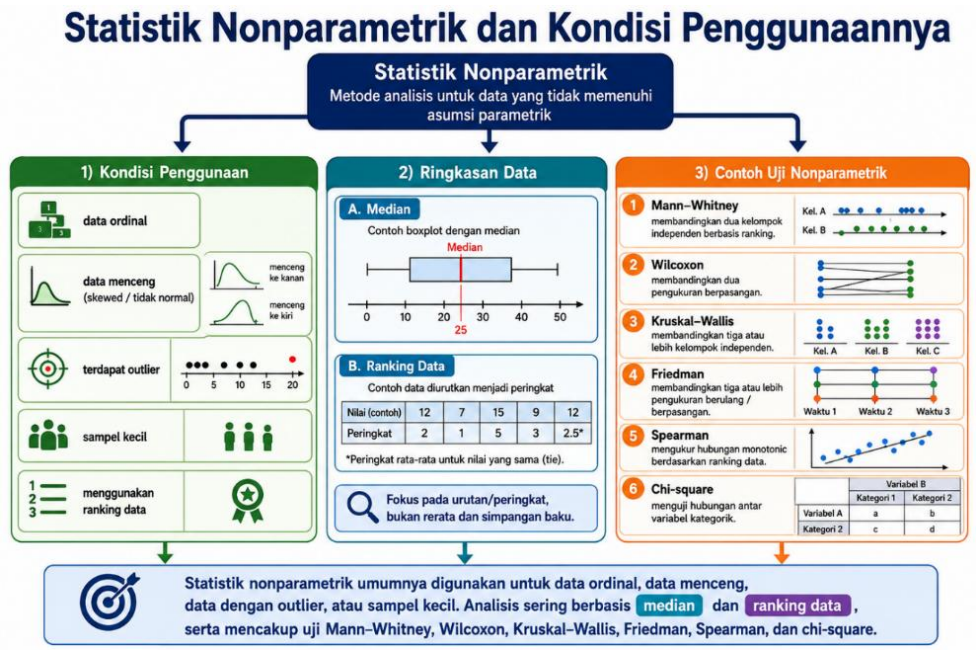
Gambar menunjukkan data numerik, distribusi normal, mean, simpangan baku, homogenitas varians, uji t, ANOVA, korelasi Pearson, dan regresi linear. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012), Chicco *et al*. (2025), serta Kaliyadan dan Kulkarni (2019).

1.8 Statistik Nonparametrik

Statistik nonparametrik adalah metode statistik yang tidak memerlukan asumsi distribusi normal secara ketat (Chicco *et al*., 2025). Statistik nonparametrik sering digunakan untuk data ordinal, data nominal, data numerik yang sangat menceng, sampel kecil, atau data dengan outlier ekstrem (Chicco *et al*., 2025). Banyak uji nonparametrik menggunakan peringkat, median, tanda, atau distribusi frekuensi sebagai dasar analisis (Chicco *et al*., 2025). Contoh uji nonparametrik adalah Mann–Whitney U test, Wilcoxon signed-rank test, Kruskal–Wallis test, Friedman test, Spearman correlation, Kendall correlation, chi-square test, dan Fisher's exact test (Chicco *et al*., 2025).

Statistik nonparametrik lebih sesuai ketika asumsi parametrik tidak terpenuhi atau ketika skala pengukuran tidak mendukung penggunaan mean sebagai ukuran utama (Chicco *et al*., 2025). Data ordinal seperti skor nyeri, tingkat kepuasan, ranking preferensi, dan skala Likert sering dianalisis dengan pendekatan nonparametrik jika asumsi interval tidak dapat dipertahankan (Kaliyadan & Kulkarni, 2019). Statistik

nonparametrik bukan metode yang kurang ilmiah karena metode ini justru lebih robust pada kondisi distribusi data yang tidak normal atau data yang bersifat ordinal (Chicco *et al*., 2025). Keterbatasan statistik nonparametrik adalah kekuatan statistiknya dapat lebih rendah dibanding uji parametrik jika asumsi parametrik sebenarnya terpenuhi (Chicco *et al*., 2025).



Gambar 1. 8. Statistik nonparametrik dan kondisi penggunaannya
Diagram menunjukkan data ordinal, data menceng, outlier, sampel kecil, ranking data, median, Mann–Whitney, Wilcoxon, Kruskal–Wallis, Friedman, Spearman, dan chi-square. **Sumber:** Diadaptasi dari Chicco *et al*. (2025) dan Kaliyadan dan Kulkarni (2019).

1.9 Perbandingan Statistik Parametrik dan Nonparametrik

Statistik parametrik dan nonparametrik merupakan dua pendekatan analisis yang dipilih berdasarkan karakteristik data dan tujuan penelitian (Chicco *et al*., 2025). Statistik parametrik lebih tepat digunakan ketika data numerik memenuhi asumsi normalitas, homogenitas varians, dan independensi observasi (Ghasemi & Zahediasl, 2012). Statistik nonparametrik lebih tepat digunakan ketika data bersifat ordinal, distribusi sangat menceng, terdapat outlier ekstrem, atau ukuran sampel kecil (Chicco *et al*., 2025). Pemilihan antara statistik parametrik dan nonparametrik harus didasarkan pada jenis variabel, desain

penelitian, distribusi data, jumlah kelompok, dan hubungan antarobservasi (Kaliyadan & Kulkarni, 2019).

Kesalahan umum dalam penelitian adalah memilih uji statistik hanya berdasarkan kebiasaan atau pilihan standar software (Rovetta, 2024). Peneliti dapat keliru menggunakan t-test untuk semua perbandingan dua kelompok tanpa memeriksa skala data dan asumsi distribusi (Chicco *et al*., 2025). Peneliti juga dapat keliru menggunakan uji nonparametrik hanya karena data tampak tidak normal tanpa mempertimbangkan ukuran sampel, outlier, dan tujuan estimasi (Ghasemi & Zahediasl, 2012). Oleh karena itu, pemilihan uji statistik harus dilakukan melalui alur keputusan yang mempertimbangkan pertanyaan penelitian, jenis data, asumsi, dan tujuan interpretasi (Chicco *et al*., 2025).

Tabel 1. 1. Perbandingan statistik parametrik dan nonparametrik

Aspek	Statistik Parametrik	Statistik Nonparametrik
Asumsi distribusi	Umumnya memerlukan normalitas atau distribusi sampling yang mendekati normal	Tidak memerlukan asumsi normalitas ketat
Skala data	Interval atau rasio	Nominal, ordinal, atau numerik tidak normal
Ukuran pemusatan utama	Mean	Median, peringkat, frekuensi
Dua kelompok independen	Independent t-test	Mann–Whitney U test
Dua kelompok berpasangan	Paired t-test	Wilcoxon signed-rank test
Lebih dari dua kelompok independen	ANOVA	Kruskal–Wallis test
Korelasi	Pearson correlation	Spearman atau Kendall correlation
Kelebihan	Lebih kuat jika asumsi terpenuhi	Lebih robust jika asumsi parametrik dilanggar
Keterbatasan	Sensitif terhadap outlier dan pelanggaran asumsi	Dapat kurang powerful jika data sebenarnya memenuhi asumsi parametrik

Sumber: Diadaptasi dari Chicco *et al*. (2025), Ghasemi dan Zahediasl (2012), serta Kaliyadan dan Kulkarni (2019).

1.10 Pemilihan Uji Statistik Berdasarkan Jenis Data dan Desain Penelitian

Pemilihan uji statistik harus dimulai dari pertanyaan penelitian yang ingin dijawab (Kaliyadan & Kulkarni, 2019). Jika tujuan analisis hanya menggambarkan karakteristik sampel, statistik deskriptif sudah memadai (Cooksey, 2020). Jika tujuan analisis adalah membandingkan dua kelompok independen dengan variabel numerik berdistribusi normal, independent t-test dapat digunakan (Chicco *et al*., 2025). Jika tujuan analisis adalah membandingkan dua kelompok berpasangan dengan data numerik yang memenuhi asumsi parametrik, paired t-test dapat digunakan (Chicco *et al*., 2025). Jika data tidak normal atau bersifat ordinal, Mann–Whitney U test atau Wilcoxon signed-rank test dapat digunakan sesuai desain penelitian (Chicco *et al*., 2025).

Jika peneliti ingin membandingkan lebih dari dua kelompok independen dengan data numerik yang memenuhi asumsi parametrik, one-way ANOVA dapat digunakan (Chicco *et al*., 2025). Jika asumsi parametrik tidak terpenuhi pada perbandingan lebih dari dua kelompok independen, Kruskal–Wallis test dapat digunakan (Chicco *et al*., 2025). Jika data kategorik dibandingkan antar kelompok, chi-square test dapat digunakan ketika syarat frekuensi terpenuhi (Chicco *et al*., 2025). Jika ukuran sampel kecil atau expected count rendah pada data kategorik, Fisher's exact test lebih sesuai digunakan (Chicco *et al*., 2025). Jika hubungan dua variabel numerik bersifat linear dan memenuhi asumsi parametrik, Pearson correlation dapat digunakan (Chicco *et al*., 2025). Jika data bersifat ordinal atau tidak normal, Spearman correlation lebih sesuai digunakan (Chicco *et al*., 2025).



Gambar 1. 9. Pohon keputusan sederhana pemilihan uji statistic

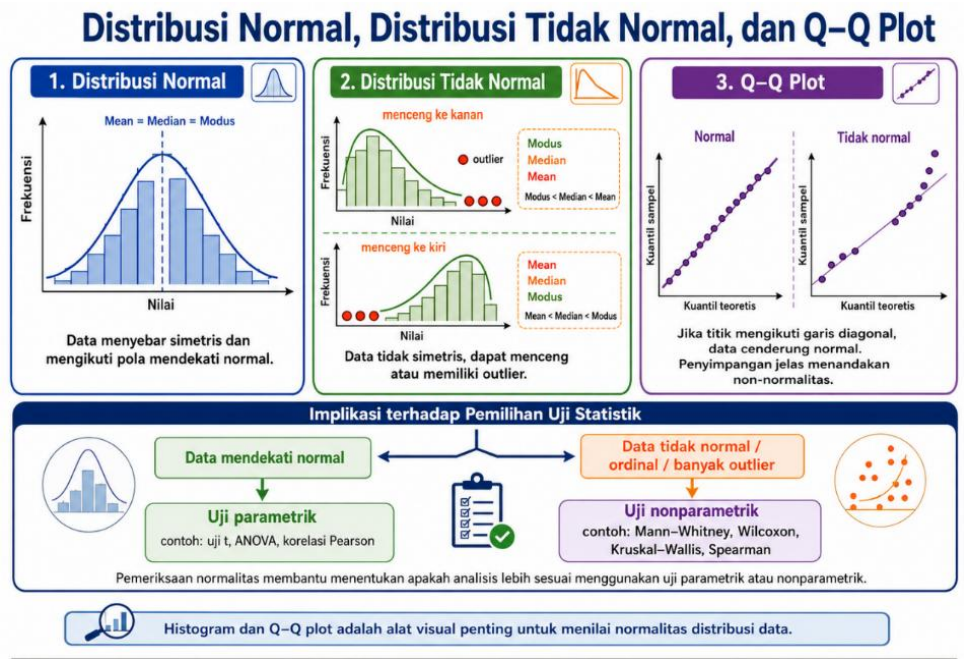
Diagram menunjukkan alur pemilihan uji berdasarkan tujuan analisis, jenis variabel, jumlah kelompok, hubungan independen atau berpasangan, normalitas data, dan skala pengukuran. **Sumber:** Diadaptasi dari Chicco *et al*. (2025), Kaliyadan dan Kulkarni (2019), serta Ghasemi dan Zahediasl (2012).

1.11 Asumsi Normalitas dan Central Limit Theorem

Normalitas adalah asumsi penting dalam banyak uji parametrik, terutama ketika ukuran sampel kecil (Ghasemi & Zahediasl, 2012). Normalitas dapat dinilai melalui histogram, Q–Q plot, Shapiro–Wilk test, Kolmogorov–Smirnov test, skewness, dan kurtosis (Ghasemi & Zahediasl, 2012). Uji normalitas memiliki keterbatasan karena sampel kecil dapat gagal mendeteksi penyimpangan distribusi yang penting (Ghasemi & Zahediasl, 2012). Uji normalitas juga dapat menjadi terlalu sensitif pada sampel sangat besar sehingga penyimpangan kecil dapat tampak signifikan secara statistik (Ghasemi & Zahediasl, 2012).

Central limit theorem menjelaskan bahwa distribusi sampling dari mean cenderung mendekati normal ketika ukuran sampel cukup besar (Ghasemi & Zahediasl, 2012). Konsep central limit theorem membantu menjelaskan mengapa beberapa uji parametrik masih dapat digunakan pada sampel besar walaupun data mentah tidak sepenuhnya normal (Ghasemi & Zahediasl, 2012). Namun, central limit theorem tidak selalu

cukup untuk data yang sangat menceng, memiliki outlier ekstrem, atau memiliki ukuran sampel kecil pada setiap kelompok (Ghasemi & Zahediasl, 2012). Peneliti tetap perlu mempertimbangkan transformasi data, uji nonparametrik, statistik robust, atau model alternatif ketika asumsi parametrik tidak sesuai (Chicco *et al*., 2025).

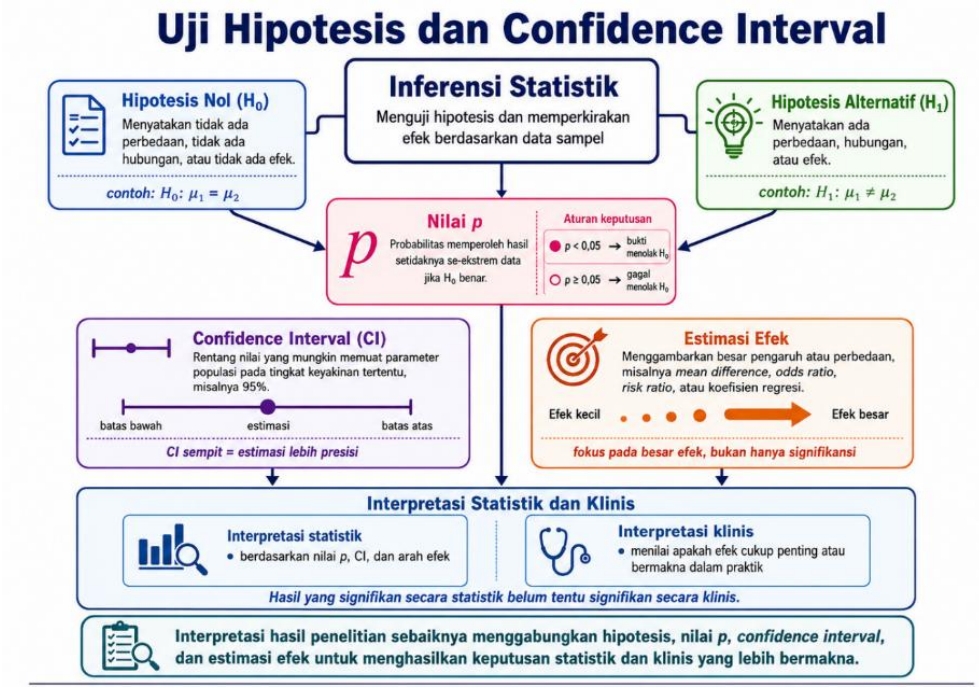


Gambar 1. 10. Distribusi normal, distribusi tidak normal, dan Q–Q plot
Gambar menunjukkan distribusi normal, distribusi menceng, outlier, histogram, Q–Q plot, dan hubungan normalitas dengan pemilihan uji parametrik atau nonparametrik. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012) serta Chicco *et al*. (2025).

1.12 Uji Hipotesis, Nilai p, dan Confidence Interval

Uji hipotesis digunakan untuk menilai apakah data sampel memberikan bukti yang cukup terhadap hipotesis nol (Rovetta, 2024). Hipotesis nol biasanya menyatakan tidak ada perbedaan, tidak ada hubungan, atau tidak ada efek (Rovetta, 2024). Nilai p menunjukkan probabilitas memperoleh hasil setidaknya se-ekstrem hasil yang diamati apabila hipotesis nol benar (Rovetta, 2024). Nilai p yang kecil menunjukkan bahwa data kurang konsisten dengan hipotesis nol, tetapi tidak secara langsung menunjukkan besar efek atau kepentingan klinis (Rovetta, 2024).

Confidence interval menunjukkan rentang nilai yang masuk akal untuk parameter populasi berdasarkan data sampel (Rovetta, 2024). Confidence interval lebih informatif daripada nilai p saja karena memberikan informasi tentang arah efek, besar efek, dan ketidakpastian estimasi (Rovetta, 2024). Confidence interval yang sempit menunjukkan estimasi yang lebih presisi dibanding confidence interval yang lebar (Rovetta, 2024). Hasil penelitian sebaiknya dilaporkan dengan estimasi efek, confidence interval, dan nilai p secara bersama-sama agar interpretasi statistik lebih lengkap (Rovetta, 2024).



Gambar 1. 11. Uji hipotesis dan confidence interval

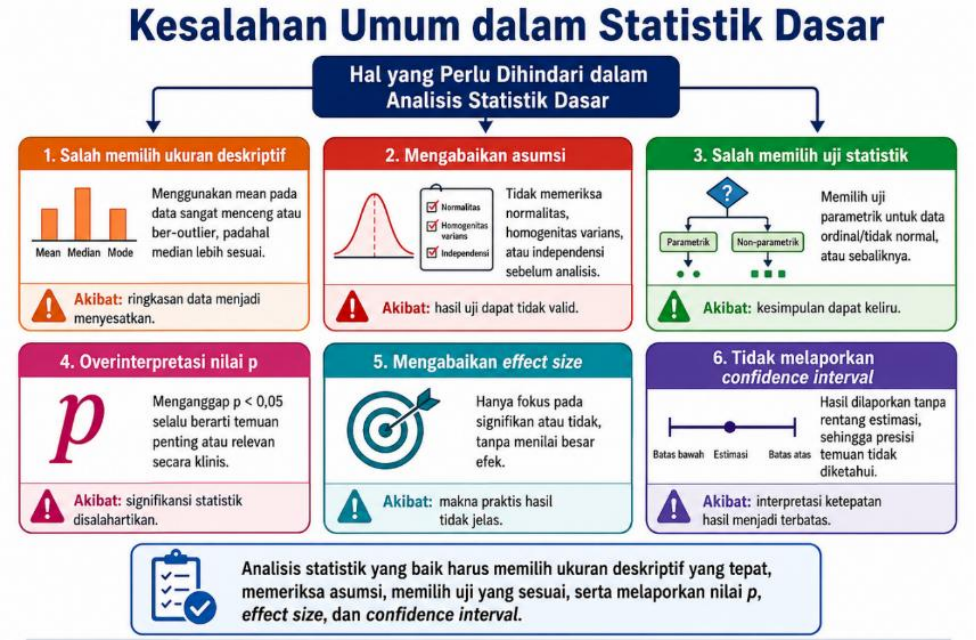
Diagram menunjukkan hipotesis nol, hipotesis alternatif, nilai p, confidence interval, estimasi efek, dan interpretasi statistik serta klinis.
Sumber: Diadaptasi dari Rovetta (2024) dan Chicco *et al*. (2025).

1.13 Kesalahan Umum dalam Penggunaan Statistik Dasar

Kesalahan umum pertama adalah menggunakan statistik deskriptif yang tidak sesuai dengan distribusi data (Cooksey, 2020). Data yang sangat menceng sebaiknya tidak hanya dilaporkan dengan mean dan simpangan baku karena median dan interquartile range dapat lebih mewakili pusat dan sebaran data (Cooksey, 2020). Kesalahan umum kedua adalah memilih uji parametrik tanpa memeriksa asumsi normalitas, homogenitas varians, dan outlier (Ghasemi & Zahediasl,

2012). Kesalahan umum ketiga adalah menggunakan uji nonparametrik tanpa mempertimbangkan skala data, desain penelitian, dan tujuan analisis (Chicco *et al*., 2025).

Kesalahan umum lain adalah menafsirkan nilai p sebagai ukuran kebenaran hipotesis atau besar efek (Rovetta, 2024). Hasil yang signifikan secara statistik belum tentu bermakna secara klinis apabila besar efeknya kecil (Rovetta, 2024). Hasil yang tidak signifikan secara statistik belum tentu membuktikan tidak adanya efek, terutama apabila ukuran sampel kecil dan confidence interval lebar (Rovetta, 2024). Peneliti perlu melaporkan effect size, confidence interval, nilai p, dan konteks ilmiah agar interpretasi hasil tidak menyesatkan (Rovetta, 2024).



Gambar 1. 12. Kesalahan umum dalam statistik dasar

Diagram menunjukkan kesalahan memilih ukuran deskriptif, mengabaikan asumsi, salah memilih uji, overinterpretasi nilai p, mengabaikan effect size, dan tidak melaporkan confidence interval. **Sumber:** Diadaptasi dari Rovetta (2024), Chicco *et al*. (2025), Cooksey (2020), serta Ghasemi dan Zahediasl (2012).

Statistik deskriptif digunakan untuk merangkum dan menyajikan data melalui frekuensi, persentase, mean, median, simpangan baku, interquartile range, histogram, boxplot, dan grafik lain (Cooksey, 2020). Statistik inferensial digunakan untuk menarik kesimpulan dari sampel ke populasi melalui estimasi, uji hipotesis, confidence interval, dan nilai p (Rovetta, 2024). Statistik parametrik digunakan ketika data dan desain

penelitian memenuhi asumsi tertentu seperti normalitas, homogenitas varians, independensi observasi, dan skala numerik yang sesuai (Ghasemi & Zahediasl, 2012). Statistik nonparametrik digunakan ketika asumsi parametrik tidak terpenuhi atau ketika data berskala ordinal, nominal, sangat menceng, atau mengandung outlier ekstrem (Chicco *et al*., 2025).

Pemilihan metode statistik harus dimulai dari pertanyaan penelitian, jenis variabel, desain penelitian, distribusi data, jumlah kelompok, dan hubungan antarobservasi (Kaliyadan & Kulkarni, 2019). Uji parametrik dapat lebih powerful bila asumsi terpenuhi, sedangkan uji nonparametrik lebih robust bila asumsi parametrik dilanggar (Chicco *et al*., 2025). Central limit theorem dapat membantu menjelaskan penggunaan uji parametrik pada sampel besar, tetapi tetap perlu disertai pemeriksaan distribusi dan outlier (Ghasemi & Zahediasl, 2012). Kompetensi statistik dasar tidak hanya berarti mampu menjalankan uji statistik, tetapi juga mampu memilih, melaporkan, dan menafsirkan hasil secara ilmiah (Rovetta, 2024).



Gambar 1. 13. Ringkasan konsep statistik dasar

Diagram merangkum statistik deskriptif, statistik inferensial, statistik parametrik, statistik nonparametrik, jenis data, asumsi normalitas, pemilihan uji, confidence interval, nilai p, dan interpretasi ilmiah. **Sumber:** Diadaptasi dari Cooksey (2020), Kaliyadan dan Kulkarni (2019), Chicco *et al*. (2025), Ghasemi dan Zahediasl (2012), serta Rovetta (2024).

Review Questions

1. **Apa perbedaan statistik deskriptif dan statistik inferensial?**
   Jawaban: Statistik deskriptif digunakan untuk merangkum dan menyajikan data yang diamati, sedangkan statistik inferensial digunakan untuk menarik kesimpulan dari sampel ke populasi.
2. **Apa contoh ukuran pemusatan data?**
   Jawaban: Contoh ukuran pemusatan data adalah mean, median, dan modus.
3. **Kapan median lebih tepat digunakan daripada mean?**
   Jawaban: Median lebih tepat digunakan ketika data menceng, memiliki outlier ekstrem, atau distribusi tidak simetris.
4. **Apa fungsi simpangan baku?**
   Jawaban: Simpangan baku menunjukkan seberapa jauh nilai observasi menyebar dari rata-rata.
5. **Apa yang dimaksud statistik inferensial?**
   Jawaban: Statistik inferensial adalah metode statistik untuk memperkirakan parameter populasi atau menguji hipotesis berdasarkan data sampel.
6. **Apa perbedaan nilai p dan confidence interval?**
   Jawaban: Nilai p menunjukkan konsistensi data terhadap hipotesis nol, sedangkan confidence interval menunjukkan rentang estimasi parameter yang masuk akal berdasarkan data sampel.
7. **Apa ciri umum statistik parametrik?**
   Jawaban: Statistik parametrik umumnya digunakan untuk data numerik dan memerlukan asumsi seperti normalitas, homogenitas varians, dan independensi observasi.
8. **Apa ciri umum statistik nonparametrik?**
   Jawaban: Statistik nonparametrik tidak memerlukan asumsi normalitas ketat dan sering digunakan untuk data ordinal, nominal, sampel kecil, atau data numerik yang sangat tidak normal.
9. **Apa contoh pasangan uji parametrik dan nonparametrik?**
   Jawaban: Independent t-test berpasangan dengan Mann–Whitney U test, paired t-test berpasangan dengan Wilcoxon signed-rank test, dan ANOVA berpasangan dengan Kruskal–Wallis test.
10. **Mengapa pemilihan uji statistik harus memperhatikan jenis data?**
    Jawaban: Pemilihan uji statistik harus memperhatikan jenis data karena setiap uji memiliki asumsi, skala pengukuran, dan desain penelitian yang sesuai.

DAFTAR PUSTAKA

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*, Article 40. https://doi.org/10.1186/s13040-025-00465-6 (PMC)

Cooksey, R. W. (2020). Descriptive statistics for summarising data. In *Illustrating statistical procedures: Finding meaning in quantitative data* (pp. 61–139). Springer.

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 (PMC)

Kaliyadan, F., & Kulkarni, V. (2019). Types of variables, descriptive statistics, and sample size. *Indian Dermatology Online Journal, 10*(1), 82–86. https://doi.org/10.4103/idoj.IDOJ_468_18 (PMC)

Rovetta, A., Mansournia, M. A., & Vitale, A. (2024). For a proper use of frequentist inferential statistics in public health and clinical research. *Global Epidemiology, 8*, 100151. https://doi.org/10.1016/j.gloepi.2024.100151 (sciencedirect.com)

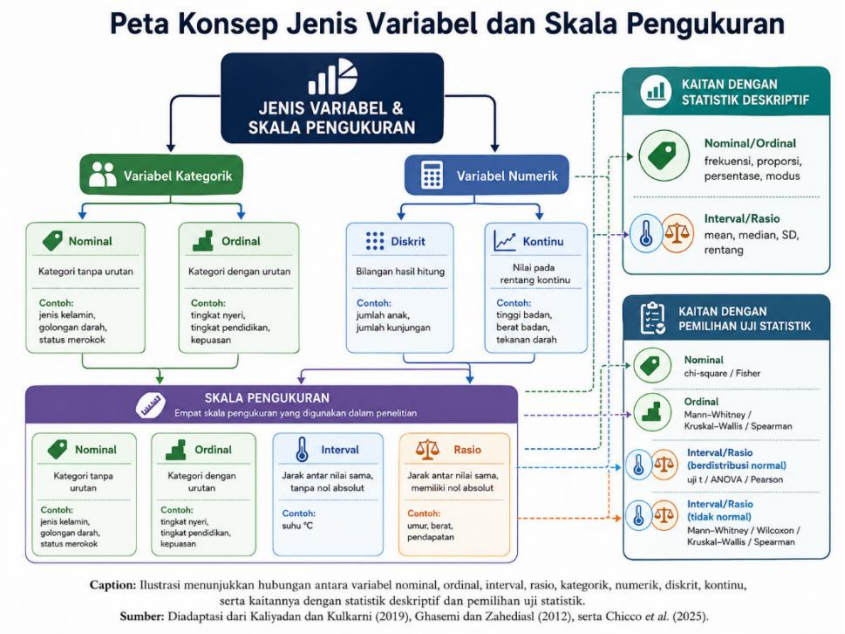
BAB 2

JENIS VARIABEL, SKALA PENGUKURAN, DATA, DAN PERSIAPAN ANALISIS STATISTIK

2.1 Pendahuluan

Pemahaman tentang jenis variabel merupakan fondasi utama sebelum peneliti memilih statistik deskriptif, statistik inferensial, uji parametrik, atau uji nonparametrik (Kaliyadan & Kulkarni, 2019). Variabel adalah karakteristik, atribut, atau pengukuran yang nilainya dapat berbeda antarindividu, objek, waktu, atau kelompok penelitian (Kaliyadan & Kulkarni, 2019). Kesalahan dalam mengidentifikasi jenis variabel dapat menyebabkan kesalahan dalam penyajian data, pemilihan uji statistik, dan interpretasi hasil penelitian (Chicco *et al*., 2025). Oleh karena itu, tahap awal analisis statistik harus dimulai dari identifikasi variabel, skala pengukuran, distribusi data, desain penelitian, dan tujuan analisis (Kaliyadan & Kulkarni, 2019; Chicco *et al*., 2025).

Dalam penelitian kuantitatif, variabel dapat diklasifikasikan berdasarkan skala pengukuran, peran dalam penelitian, dan karakteristik datanya (Kaliyadan & Kulkarni, 2019). Berdasarkan skala pengukuran, variabel dapat berupa nominal, ordinal, interval, atau rasio (Kaliyadan & Kulkarni, 2019). Berdasarkan perannya, variabel dapat berupa variabel independen, dependen, kovariat, confounder, moderator, mediator, atau variabel kontrol (Kaliyadan & Kulkarni, 2019). Berdasarkan karakteristik datanya, variabel dapat berupa kategorik, numerik, diskrit, kontinu, terdistribusi normal, tidak normal, atau mengandung outlier (Ghasemi & Zahediasl, 2012).



Caption: Ilustrasi menunjukkan hubungan antara variabel nominal, ordinal, interval, rasio, kategorik, numerik, diskrit, kontinu, serta kaitannya dengan statistik deskriptif dan pemilihan uji statistik.
Sumber: Diadaptasi dari Kaliyadan dan Kulkarni (2019), Ghasemi dan Zahediasl (2012), serta Chicco *et al.* (2025).

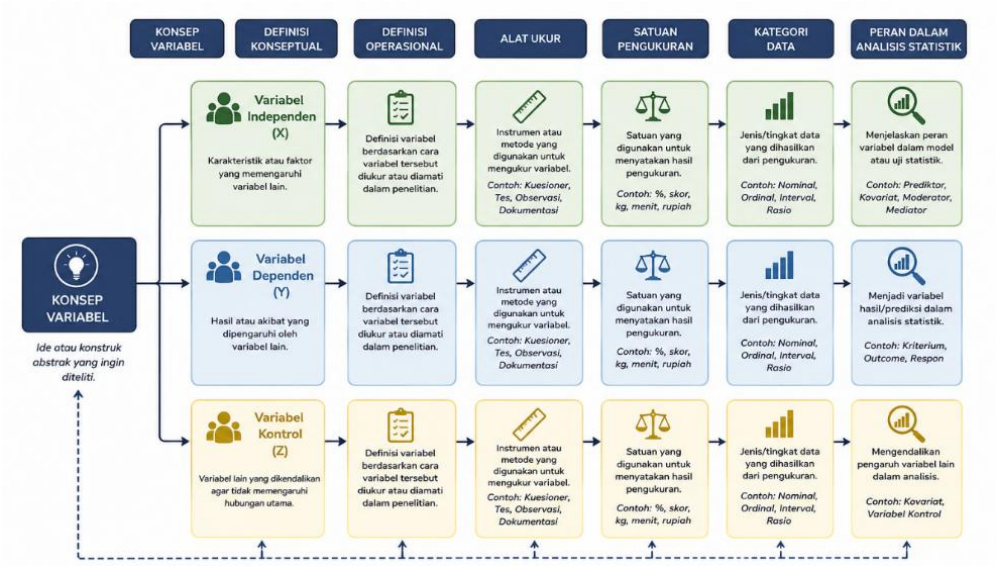
Gambar 2. 1. Peta konsep jenis variabel dan skala pengukuran
Ilustrasi menunjukkan hubungan antara variabel nominal, ordinal, interval, rasio, kategorik, numerik, diskrit, kontinu, serta kaitannya dengan statistik deskriptif dan pemilihan uji statistik. **Sumber:** Diadaptasi dari Kaliyadan dan Kulkarni (2019), Ghasemi dan Zahediasl (2012), serta Chicco *et al*. (2025).

2.2 Definisi Variabel dalam Penelitian

Variabel adalah atribut yang dapat diukur, diamati, atau dikategorikan dalam suatu penelitian (Kaliyadan & Kulkarni, 2019). Variabel dapat berupa karakteristik demografis seperti usia, jenis kelamin, pendidikan, dan pekerjaan (Kaliyadan & Kulkarni, 2019). Variabel juga dapat berupa data klinis seperti tekanan darah, kadar glukosa, indeks massa tubuh, skor nyeri, atau diagnosis penyakit (Kaliyadan & Kulkarni, 2019). Dalam penelitian biomedis dan kesehatan, variabel menjadi dasar untuk menggambarkan populasi, membandingkan kelompok, menguji hubungan, dan membangun model prediksi (Chicco *et al*., 2025).

Variabel harus didefinisikan secara operasional agar pengukuran dapat dilakukan secara konsisten (Kaliyadan & Kulkarni, 2019). Definisi operasional menjelaskan bagaimana variabel diukur, alat ukur apa yang digunakan, satuan apa yang dipakai, dan bagaimana hasilnya dikategorikan (Kaliyadan & Kulkarni, 2019). Misalnya, variabel tekanan darah dapat didefinisikan sebagai tekanan sistolik dan diastolik yang

diukur dengan sphygmomanometer dalam satuan mmHg (Kaliyadan & Kulkarni, 2019). Definisi operasional yang buruk dapat menyebabkan bias pengukuran dan menurunkan validitas hasil penelitian (Rovetta, 2024).



Gambar 2. 2. Definisi variabel dan operasionalisasi variabel

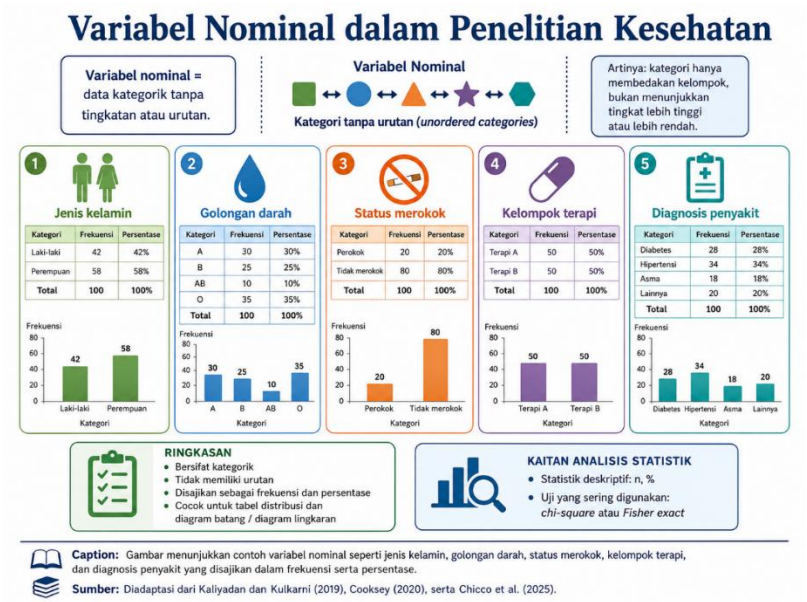
Diagram menunjukkan konsep variabel, definisi konseptual, definisi operasional, alat ukur, satuan pengukuran, kategori data, dan peran variabel dalam analisis statistik. **Sumber:** Diadaptasi dari Kaliyadan dan Kulkarni (2019) serta Rovetta (2024).

2.3 Variabel Nominal

Variabel nominal adalah variabel kategorik yang membedakan kelompok tanpa urutan atau tingkatan alami (Kaliyadan & Kulkarni, 2019). Contoh variabel nominal adalah jenis kelamin, golongan darah, status merokok, jenis pekerjaan, kelompok terapi, dan diagnosis penyakit (Kaliyadan & Kulkarni, 2019). Pada variabel nominal, angka yang diberikan hanya berfungsi sebagai label dan tidak menunjukkan nilai kuantitatif (Kaliyadan & Kulkarni, 2019). Misalnya, kode 1 untuk laki-laki dan kode 2 untuk perempuan tidak berarti perempuan memiliki nilai lebih tinggi daripada laki-laki (Kaliyadan & Kulkarni, 2019).

Statistik deskriptif untuk variabel nominal biasanya berupa frekuensi dan persentase (Kaliyadan & Kulkarni, 2019). Visualisasi yang sesuai untuk variabel nominal adalah bar chart atau pie chart, meskipun bar chart sering lebih mudah dibaca untuk perbandingan kategori (Cooksey, 2020). Uji statistik untuk variabel nominal biasanya mencakup chi-square test atau Fisher's exact test ketika peneliti ingin

membandingkan proporsi antar kelompok (Chicco *et al*., 2025). Fisher's exact test lebih sesuai digunakan ketika ukuran sampel kecil atau expected count dalam tabel kontingensi rendah (Chicco *et al*., 2025).



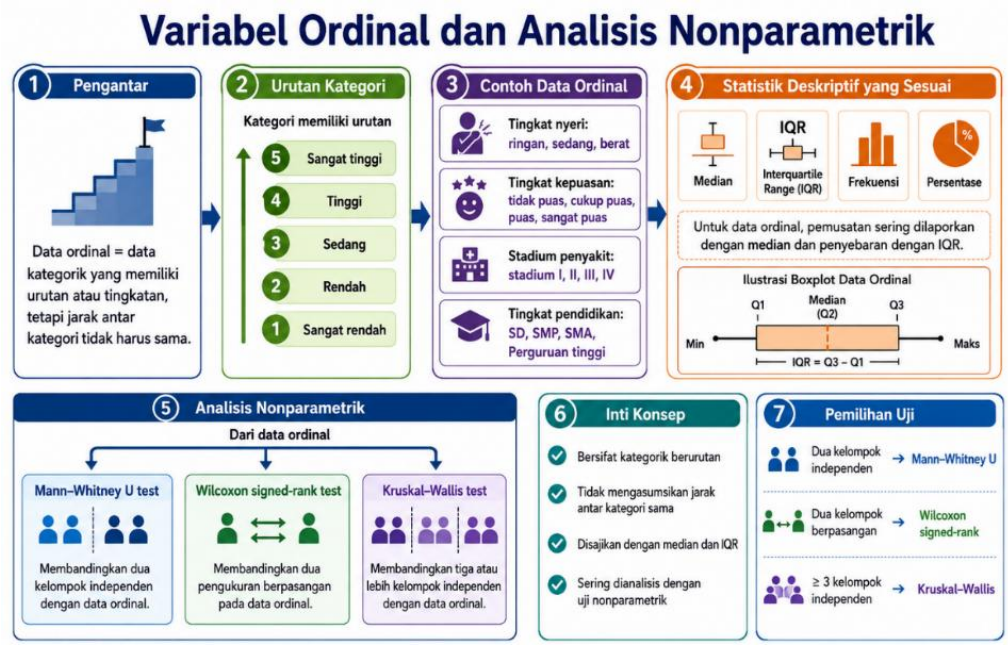
Gambar 2. 3. Variabel nominal dalam penelitian Kesehatan
Gambar menunjukkan contoh variabel nominal seperti jenis kelamin, golongan darah, status merokok, kelompok terapi, dan diagnosis penyakit yang disajikan dalam frekuensi serta persentase. **Sumber:** Diadaptasi dari Kaliyadan dan Kulkarni (2019), Cooksey (2020), serta Chicco *et al*. (2025).

2.4 Variabel Ordinal

Variabel ordinal adalah variabel kategorik yang memiliki urutan atau tingkatan, tetapi jarak antar kategori tidak dapat dianggap sama secara pasti (Kaliyadan & Kulkarni, 2019). Contoh variabel ordinal adalah tingkat nyeri ringan-sedang-berat, stadium penyakit, tingkat pendidikan, derajat kepuasan, dan skor Likert (Kaliyadan & Kulkarni, 2019). Pada variabel ordinal, urutan kategori memiliki makna, tetapi selisih antara kategori pertama dan kedua belum tentu sama dengan selisih antara kategori kedua dan ketiga (Kaliyadan & Kulkarni, 2019). Oleh karena itu, analisis ordinal harus mempertimbangkan sifat berurutan tetapi tidak selalu berskala interval (Chicco *et al*., 2025).

Statistik deskriptif untuk variabel ordinal dapat berupa frekuensi, persentase, median, dan interquartile range (Kaliyadan & Kulkarni, 2019). Uji nonparametrik sering digunakan pada data ordinal karena data ordinal tidak selalu memenuhi asumsi interval dan normalitas (Chicco *et*

al., 2025). Mann–Whitney U test dapat digunakan untuk membandingkan dua kelompok independen dengan data ordinal (Chicco *et al*., 2025). Wilcoxon signed-rank test dapat digunakan untuk membandingkan dua pengukuran berpasangan pada data ordinal atau data numerik yang tidak normal (Chicco *et al*., 2025). Kruskal–Wallis test dapat digunakan untuk membandingkan lebih dari dua kelompok independen pada data ordinal atau data numerik yang tidak memenuhi asumsi parametrik (Chicco *et al*., 2025).



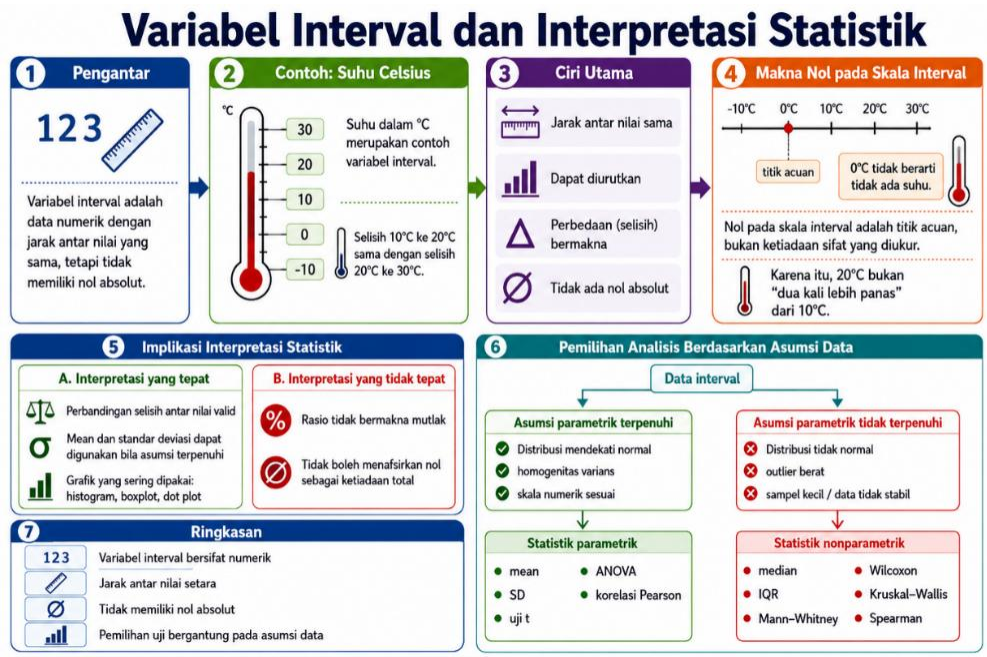
Gambar 2. 4. Variabel ordinal dan analisis nonparametric
Diagram menunjukkan data ordinal, urutan kategori, median, interquartile range, Mann–Whitney U test, Wilcoxon signed-rank test, dan Kruskal–Wallis test.
Sumber: Diadaptasi dari Kaliyadan dan Kulkarni (2019) serta Chicco *et al*. (2025).

2.5 Variabel Interval

Variabel interval adalah variabel numerik yang memiliki jarak antar nilai yang sama, tetapi tidak memiliki nol absolut (Kaliyadan & Kulkarni, 2019). Contoh variabel interval adalah suhu dalam Celsius atau Fahrenheit karena nilai nol tidak menunjukkan ketiadaan suhu (Kaliyadan & Kulkarni, 2019). Pada skala interval, operasi penjumlahan dan pengurangan bermakna, tetapi rasio seperti “dua kali lebih besar”

tidak selalu bermakna secara konseptual (Kaliyadan & Kulkarni, 2019). Skala interval dapat dianalisis dengan metode parametrik bila data memenuhi asumsi distribusi dan desain penelitian yang sesuai (Ghasemi & Zahediasl, 2012).

Statistik deskriptif untuk variabel interval dapat berupa mean dan simpangan baku bila distribusi data normal (Ghasemi & Zahediasl, 2012). Median dan interquartile range dapat digunakan bila distribusi data menceng atau terdapat outlier (Cooksey, 2020). Uji parametrik seperti t-test, ANOVA, Pearson correlation, dan linear regression dapat digunakan pada data interval bila asumsi statistik terpenuhi (Chicco *et al*., 2025). Jika asumsi parametrik tidak terpenuhi, peneliti dapat mempertimbangkan uji nonparametrik, transformasi data, atau pendekatan robust sesuai tujuan analisis (Chicco *et al*., 2025).



Gambar 2. 5. Variabel interval dan interpretasi statistic

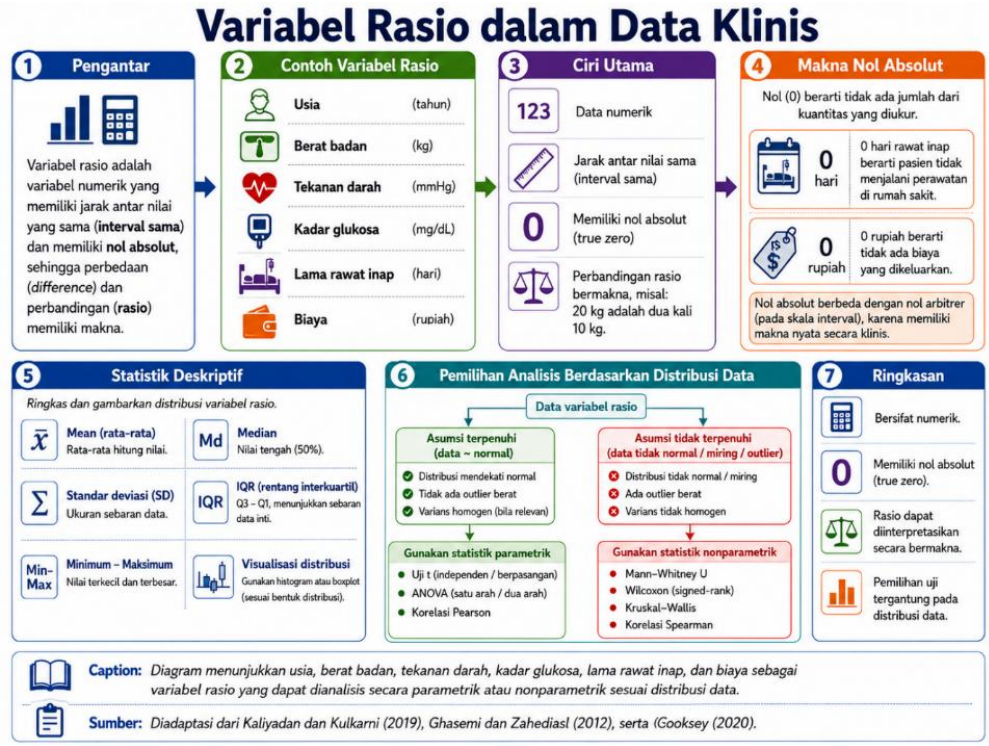
Gambar menunjukkan contoh suhu Celsius sebagai variabel interval, jarak antar nilai yang sama, tidak adanya nol absolut, serta pemilihan statistik parametrik atau nonparametrik berdasarkan asumsi data.

Sumber: Diadaptasi dari Kaliyadan dan Kulkarni (2019), Ghasemi dan Zahediasl (2012), serta Chicco *et al*. (2025).

2.6 Variabel Rasio

Variabel rasio adalah variabel numerik yang memiliki jarak antar nilai yang sama dan memiliki nol absolut (Kaliyadan & Kulkarni, 2019). Contoh variabel rasio adalah usia, berat badan, tinggi badan, kadar glukosa, tekanan darah, lama rawat inap, pendapatan, dan jumlah kunjungan klinik (Kaliyadan & Kulkarni, 2019). Nol absolut berarti nilai nol menunjukkan ketiadaan kuantitas yang diukur (Kaliyadan & Kulkarni, 2019). Karena memiliki nol absolut, pernyataan rasio seperti "dua kali lebih tinggi" dapat bermakna pada variabel rasio tertentu (Kaliyadan & Kulkarni, 2019).

Variabel rasio dapat dianalisis dengan statistik parametrik bila distribusi data dan asumsi analisis terpenuhi (Ghasemi & Zahediasl, 2012). Variabel rasio yang berdistribusi normal dapat diringkas dengan mean dan simpangan baku (Ghasemi & Zahediasl, 2012). Variabel rasio yang menceng dapat diringkas dengan median dan interquartile range (Cooksey, 2020). Lama rawat inap dan biaya pelayanan kesehatan sering menceng ke kanan sehingga median dan interquartile range dapat lebih representatif daripada mean dan simpangan baku (Cooksey, 2020).



Gambar 2. 6. Variabel rasio dalam data klinis

Diagram menunjukkan usia, berat badan, tekanan darah, kadar glukosa, lama rawat inap, dan biaya sebagai variabel rasio yang dapat dianalisis secara parametrik atau nonparametrik sesuai distribusi data.
Diadaptasi dari Kaliyadan dan Kulkarni (2019), Ghasemi dan Zahediasl (2012), serta Cooksey (2020).

2.7 Variabel Kategorik dan Variabel Numerik

Variabel kategorik adalah variabel yang nilainya berbentuk kategori atau kelompok (Kaliyadan & Kulkarni, 2019). Variabel kategorik dapat berupa nominal atau ordinal (Kaliyadan & Kulkarni, 2019). Contoh variabel kategorik nominal adalah jenis kelamin, golongan darah, dan status terapi (Kaliyadan & Kulkarni, 2019). Contoh variabel kategorik ordinal adalah derajat nyeri, stadium penyakit, dan tingkat kepuasan pasien (Kaliyadan & Kulkarni, 2019).

Variabel numerik adalah variabel yang nilainya berupa angka yang memiliki makna kuantitatif (Kaliyadan & Kulkarni, 2019). Variabel numerik dapat berupa interval atau rasio (Kaliyadan & Kulkarni, 2019). Contoh variabel numerik adalah usia, berat badan, kadar hemoglobin, tekanan darah, skor laboratorium, dan lama rawat inap (Kaliyadan & Kulkarni, 2019). Variabel numerik dapat dianalisis dengan metode

parametrik atau nonparametrik bergantung pada distribusi data, ukuran sampel, dan asumsi analisis (Ghasemi & Zahediasl, 2012; Chicco *et al*., 2025).

Tabel 2. 1. Perbandingan variabel kategorik dan numerik

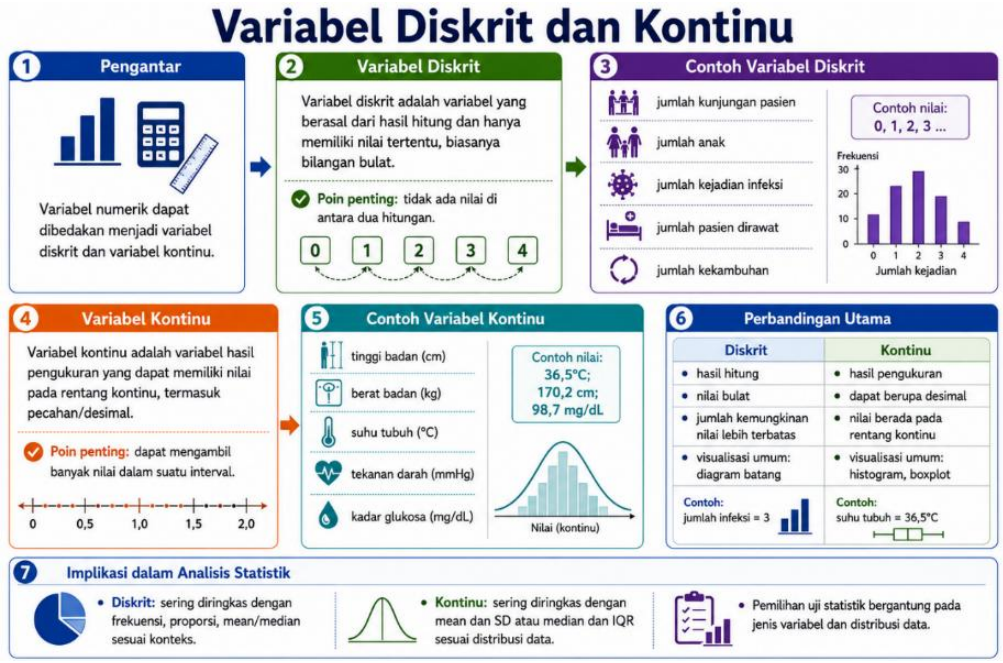
Aspek	Variabel Kategorik	Variabel Numerik
Bentuk data	Kategori atau kelompok	Angka kuantitatif
Subtipe	Nominal dan ordinal	Interval dan rasio
Contoh	Jenis kelamin, stadium penyakit, status merokok	Usia, tekanan darah, kadar glukosa
Ringkasan umum	Frekuensi dan persentase	Mean-SD atau median-IQR
Uji umum	Chi-square, Fisher's exact, uji ordinal	t-test, ANOVA, korelasi, regresi, atau uji nonparametrik

Sumber: Diadaptasi dari Kaliyadan dan Kulkarni (2019), Ghasemi dan Zahediasl (2012), serta Chicco *et al*. (2025).

2.8 Variabel Diskrit dan Kontinu

Variabel diskrit adalah variabel numerik yang nilainya berupa hitungan dan biasanya tidak memiliki nilai pecahan bermakna di antara dua angka bulat (Kaliyadan & Kulkarni, 2019). Contoh variabel diskrit adalah jumlah anak, jumlah kunjungan pasien, jumlah episode infeksi, jumlah gigi hilang, dan jumlah kejadian efek samping (Kaliyadan & Kulkarni, 2019). Variabel diskrit sering dianalisis sebagai count data dan dapat memerlukan pendekatan statistik khusus bila distribusinya tidak normal (Kaliyadan & Kulkarni, 2019). Jumlah kejadian yang jarang sering tidak cocok dianalisis dengan asumsi normal sederhana karena distribusinya dapat menceng atau mengikuti pola Poisson (Kaliyadan & Kulkarni, 2019).

Variabel kontinu adalah variabel numerik yang secara teoritis dapat memiliki nilai pada rentang kontinu (Kaliyadan & Kulkarni, 2019). Contoh variabel kontinu adalah tinggi badan, berat badan, kadar kolesterol, tekanan darah, kadar glukosa, dan waktu reaksi (Kaliyadan & Kulkarni, 2019). Variabel kontinu sering dianalisis dengan statistik parametrik bila memenuhi asumsi normalitas dan homogenitas varians (Ghasemi & Zahediasl, 2012). Jika variabel kontinu menceng atau memiliki outlier kuat, median, interquartile range, transformasi data, atau uji nonparametrik dapat dipertimbangkan (Ghasemi & Zahediasl, 2012; Chicco *et al*., 2025).



Gambar 2. 7. Variabel diskrit dan kontinu

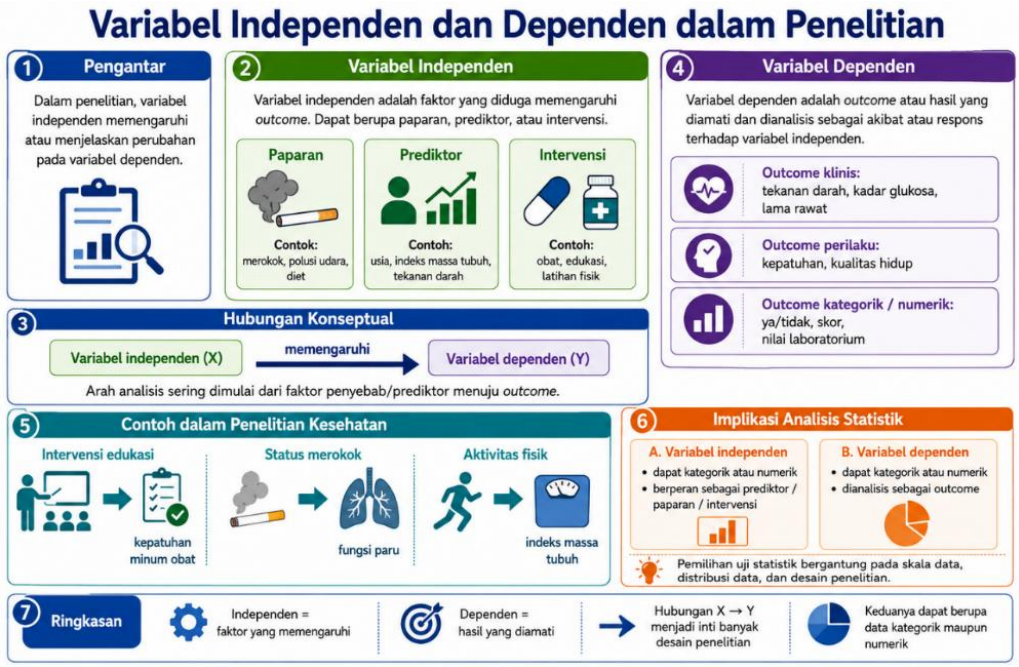
Gambar menunjukkan variabel diskrit berupa jumlah kejadian dan variabel kontinu berupa pengukuran biologis yang dapat memiliki nilai pada rentang kontinu. **Sumber:** Diadaptasi dari Kaliyadan dan Kulkarni (2019), Ghasemi dan Zahediasl (2012), serta Chicco *et al*. (2025).

2.9 Variabel Independen dan Dependen

Variabel independen adalah variabel yang dianggap sebagai faktor penjelas, paparan, prediktor, atau intervensi dalam penelitian (Kaliyadan & Kulkarni, 2019). Variabel dependen adalah variabel hasil atau outcome yang ingin dijelaskan, diprediksi, atau dibandingkan antar kelompok (Kaliyadan & Kulkarni, 2019). Dalam penelitian klinis, variabel independen dapat berupa jenis terapi, status merokok, kelompok usia, atau kadar biomarker (Kaliyadan & Kulkarni, 2019). Variabel dependen dapat berupa tekanan darah, kejadian penyakit, tingkat nyeri, lama rawat inap, atau kualitas hidup (Kaliyadan & Kulkarni, 2019).

Hubungan antara variabel independen dan dependen harus ditentukan berdasarkan pertanyaan penelitian dan desain studi (Rovetta, 2024). Dalam uji perbandingan kelompok, variabel independen sering berupa kelompok penelitian dan variabel dependen berupa outcome yang dibandingkan (Chicco *et al*., 2025). Dalam analisis regresi, variabel independen digunakan untuk menjelaskan variasi variabel dependen (Chicco *et al*., 2025). Penentuan variabel independen dan dependen yang

tidak jelas dapat menyebabkan kesalahan dalam pemilihan uji statistik dan interpretasi hasil (Rovetta, 2024).



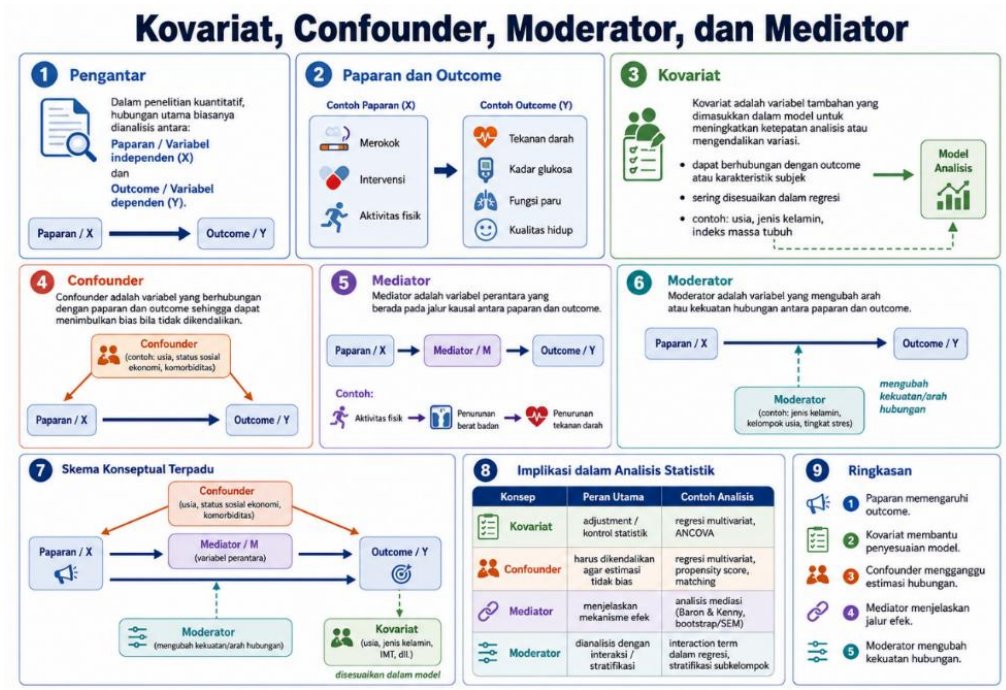
Gambar 2. 8. Variabel independen dan dependen dalam penelitian
Diagram menunjukkan variabel independen sebagai paparan, prediktor, atau intervensi dan variabel dependen sebagai outcome yang dianalisis.
Sumber: Diadaptasi dari Kaliyadan dan Kulkarni (2019), Rovetta (2024), serta Chicco *et al*. (2025).

2.10 Kovariat, Confounder, Moderator, dan Mediator

Kovariat adalah variabel tambahan yang dimasukkan dalam analisis untuk meningkatkan ketepatan estimasi atau mengendalikan perbedaan karakteristik antar kelompok (Rovetta, 2024). Confounder adalah variabel yang berhubungan dengan paparan dan outcome sehingga dapat mengaburkan hubungan sebenarnya antara variabel independen dan dependen (Rovetta, 2024). Contoh confounder dalam penelitian klinis adalah usia, jenis kelamin, status merokok, komorbiditas, dan tingkat keparahan penyakit (Rovetta, 2024). Jika confounder tidak dikendalikan, estimasi efek dapat menjadi bias (Rovetta, 2024).

Moderator adalah variabel yang memengaruhi arah atau kekuatan hubungan antara variabel independen dan dependen (Rovetta, 2024). Mediator adalah variabel yang menjelaskan mekanisme hubungan antara variabel independen dan dependen (Rovetta, 2024). Misalnya, hubungan

antara intervensi olahraga dan tekanan darah dapat dimediasi oleh penurunan berat badan (Rovetta, 2024). Pembedaan antara confounder, moderator, dan mediator penting karena masing-masing membutuhkan strategi analisis dan interpretasi yang berbeda (Rovetta, 2024).



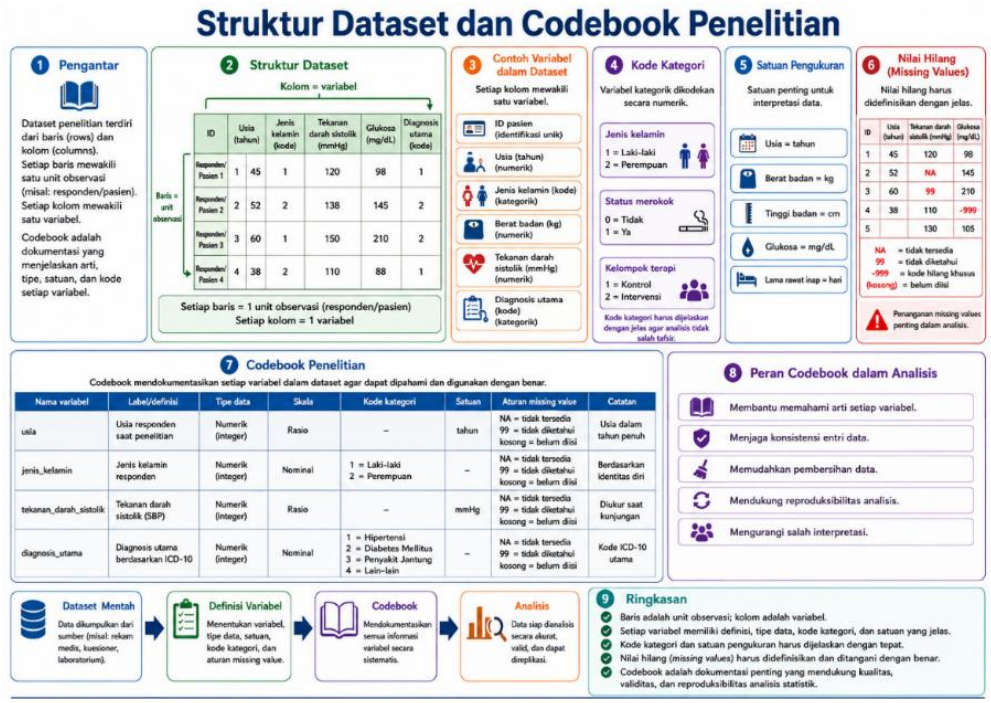
Gambar 2. 9. Kovariat, confounder, moderator, dan mediator

Gambar menunjukkan hubungan antara paparan, outcome, kovariat, confounder, moderator, dan mediator dalam rancangan analisis statistik.
Sumber: Diadaptasi dari Rovetta (2024) dan prinsip epidemiologi kuantitatif.

2.11 Coding Data dan Struktur Dataset

Coding data adalah proses mengubah informasi penelitian menjadi bentuk numerik atau kategorik yang dapat dianalisis dalam perangkat lunak statistik (Kaliyadan & Kulkarni, 2019). Variabel nominal dapat diberi kode angka sebagai label, tetapi angka tersebut tidak boleh ditafsirkan sebagai nilai kuantitatif (Kaliyadan & Kulkarni, 2019). Variabel ordinal dapat diberi kode berdasarkan urutan kategori, tetapi jarak antar kode tidak selalu mewakili jarak yang sama secara matematis (Kaliyadan & Kulkarni, 2019). Variabel numerik sebaiknya dicatat dalam satuan asli yang jelas agar analisis dan interpretasi lebih akurat (Kaliyadan & Kulkarni, 2019).

Struktur dataset yang baik biasanya menempatkan setiap baris sebagai satu unit observasi dan setiap kolom sebagai satu variabel (Cooksey, 2020). Nama variabel sebaiknya konsisten, singkat, informatif, dan tidak ambigu (Cooksey, 2020). Dataset sebaiknya dilengkapi dengan codebook yang menjelaskan nama variabel, definisi operasional, jenis data, satuan, kode kategori, nilai hilang, dan aturan pengukuran (Cooksey, 2020). Codebook yang baik membantu mencegah kesalahan analisis dan meningkatkan reprodusibilitas penelitian (Rovetta, 2024).



Gambar 2. 10. Struktur dataset dan codebook penelitian

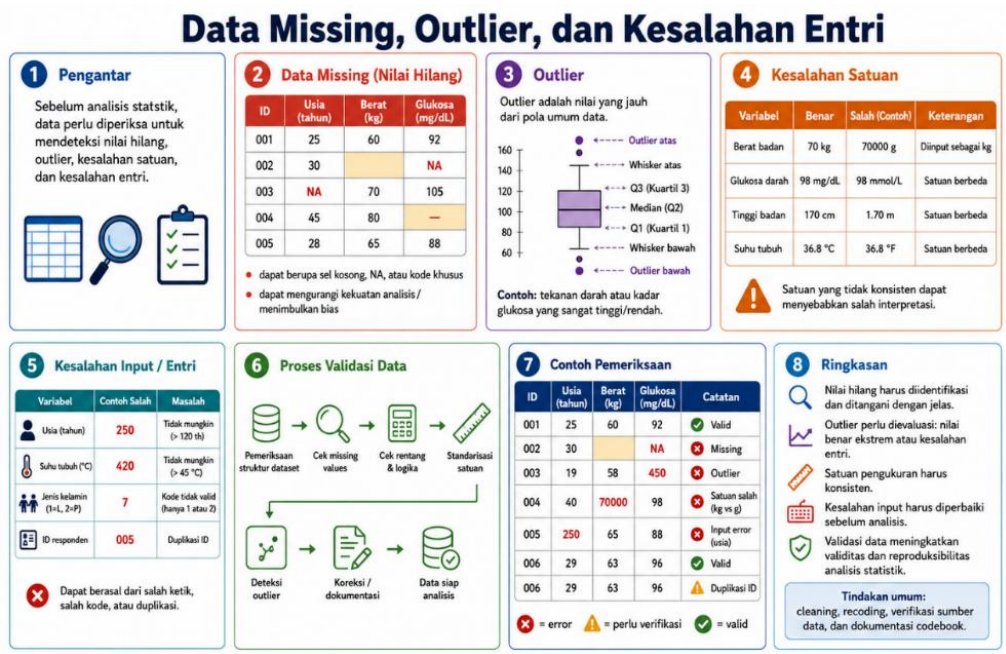
Diagram menunjukkan baris sebagai unit observasi, kolom sebagai variabel, kode kategori, satuan pengukuran, nilai hilang, dan codebook sebagai dokumentasi analisis. **Sumber:** Diadaptasi dari Cooksey (2020), Kaliyadan dan Kulkarni (2019), serta Rovetta (2024).

2.12 Data Missing, Outlier, dan Kesalahan Entri

Data missing adalah kondisi ketika nilai suatu variabel tidak tersedia untuk satu atau lebih observasi (Rovetta, 2024). Data missing dapat terjadi karena responden tidak menjawab, alat ukur gagal, data hilang saat pencatatan, atau kriteria pemeriksaan tidak terpenuhi (Rovetta, 2024). Penanganan data missing harus dilaporkan secara transparan karena dapat memengaruhi validitas hasil penelitian (Rovetta,

2024). Menghapus seluruh observasi dengan data missing dapat mengurangi ukuran sampel dan menimbulkan bias bila missing tidak terjadi secara acak (Rovetta, 2024).

Outlier adalah nilai yang sangat berbeda dari pola umum data (Ghasemi & Zahediasl, 2012). Outlier dapat mencerminkan nilai biologis yang benar, kesalahan entri, kesalahan alat ukur, atau unit pengukuran yang tidak konsisten (Ghasemi & Zahediasl, 2012). Outlier harus diperiksa sebelum analisis karena dapat memengaruhi mean, simpangan baku, korelasi, regresi, dan uji parametrik (Ghasemi & Zahediasl, 2012). Boxplot, histogram, scatterplot, dan pemeriksaan data mentah dapat membantu mendeteksi outlier dan kesalahan entri (Cooksey, 2020).



Gambar 2. 11. Data missing, outlier, dan kesalahan entri

Gambar menunjukkan nilai hilang, outlier pada boxplot, kesalahan satuan, kesalahan input, dan proses validasi data sebelum analisis statistik.

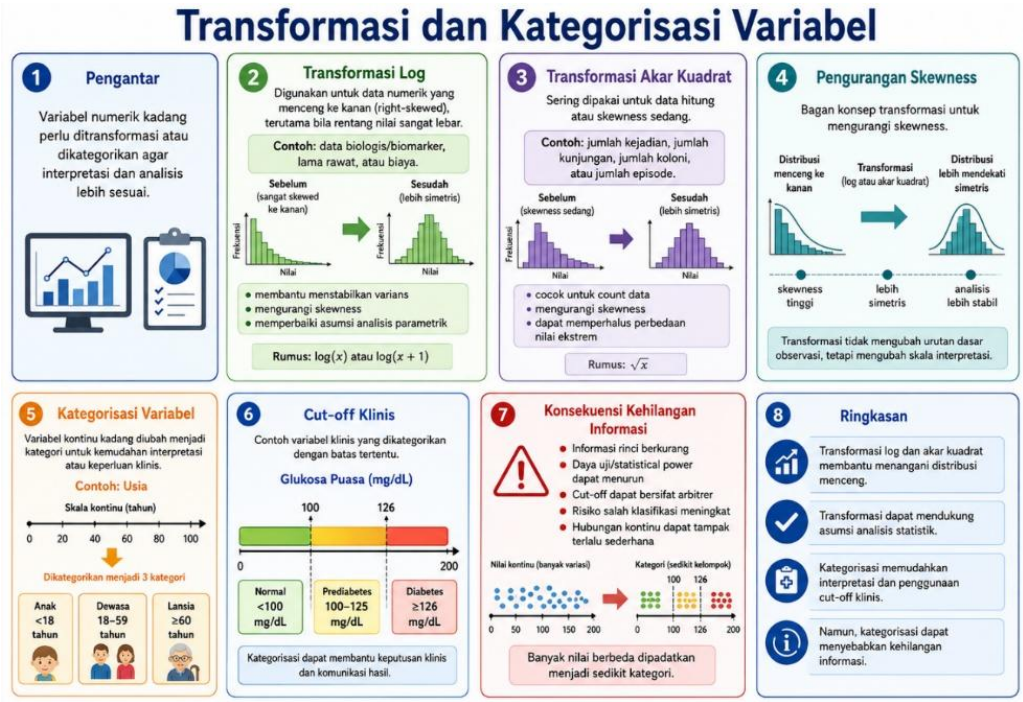
Sumber: Diadaptasi dari Ghasemi dan Zahediasl (2012), Cooksey (2020), serta Rovetta (2024).

2.13 Transformasi Data dan Pengelompokan Variabel

Transformasi data adalah proses mengubah skala variabel untuk memperbaiki distribusi, mengurangi kemencengan, menstabilkan varians, atau memenuhi asumsi model statistik (Ghasemi & Zahediasl, 2012). Transformasi yang sering digunakan adalah transformasi

logaritmik, akar kuadrat, dan reciprocal pada data yang sangat menceng ke kanan (Ghasemi & Zahediasl, 2012). Transformasi data harus dilakukan dengan alasan metodologis yang jelas dan dilaporkan secara transparan dalam metode penelitian (Rovetta, 2024). Interpretasi hasil setelah transformasi harus memperhatikan skala baru yang digunakan dalam analisis (Rovetta, 2024).

Pengelompokan variabel numerik menjadi kategori dapat membantu interpretasi klinis, tetapi dapat mengurangi informasi dan kekuatan statistik (Rovetta, 2024). Misalnya, usia dapat dikategorikan menjadi kelompok muda, dewasa, dan lansia, tetapi pengelompokan tersebut menghilangkan variasi usia yang lebih rinci (Rovetta, 2024). Kategorisasi variabel kontinu sebaiknya dilakukan berdasarkan dasar klinis, teori, guideline, atau cut-off yang telah ditentukan sebelum analisis (Rovetta, 2024). Kategorisasi yang dilakukan setelah melihat data dapat meningkatkan risiko bias interpretasi (Rovetta, 2024).



Gambar 2. 12. Transformasi dan kategorisasi variable

Diagram menunjukkan transformasi log, transformasi akar kuadrat, pengurangan skewness, kategorisasi usia, cut-off klinis, serta konsekuensi kehilangan informasi. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012) serta Rovetta (2024).

2.14 Hubungan Jenis Variabel dengan Statistik Deskriptif

Jenis variabel menentukan cara penyajian statistik deskriptif yang tepat (Kaliyadan & Kulkarni, 2019). Variabel nominal sebaiknya disajikan dengan frekuensi dan persentase (Kaliyadan & Kulkarni, 2019). Variabel ordinal dapat disajikan dengan frekuensi, persentase, median, dan interquartile range (Kaliyadan & Kulkarni, 2019). Variabel numerik yang berdistribusi normal dapat disajikan dengan mean dan simpangan baku (Ghasemi & Zahediasl, 2012). Variabel numerik yang tidak normal dapat disajikan dengan median dan interquartile range (Ghasemi & Zahediasl, 2012).

Penyajian statistik deskriptif yang tidak sesuai dapat menimbulkan interpretasi yang keliru (Cooksey, 2020). Data kategorik tidak tepat diringkas dengan mean karena angka pada kategori biasanya hanya label (Kaliyadan & Kulkarni, 2019). Data ordinal tidak selalu tepat diringkas dengan mean bila jarak antar kategori tidak dapat dianggap sama (Kaliyadan & Kulkarni, 2019). Data numerik dengan outlier ekstrem tidak tepat hanya diringkas dengan mean karena mean sensitif terhadap nilai ekstrem (Ghasemi & Zahediasl, 2012). Oleh karena itu, peneliti perlu memeriksa skala data dan distribusi sebelum membuat tabel karakteristik sampel (Cooksey, 2020).

Tabel 2.2. Hubungan jenis variabel dengan statistik deskriptif

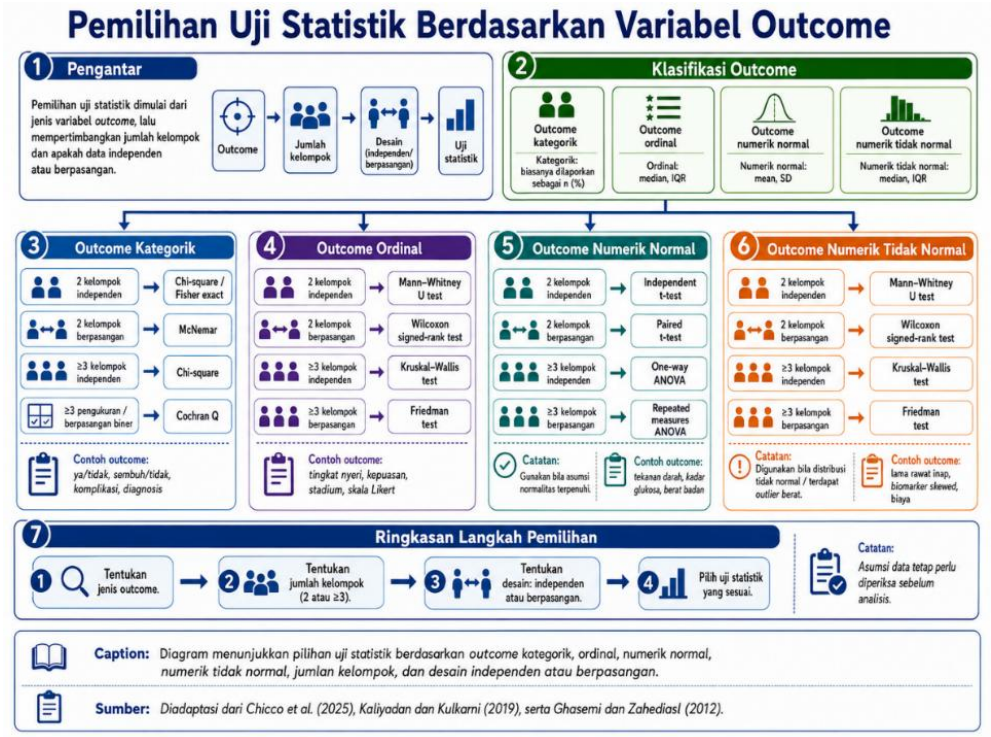
Jenis Variabel	Contoh	Statistik Deskriptif Utama	Visualisasi Umum
Nominal	Jenis kelamin, golongan darah	Frekuensi, persentase	Bar chart
Ordinal	Skor nyeri, tingkat kepuasan	Frekuensi, persentase, median, IQR	Bar chart, stacked bar
Numerik normal	Tekanan darah normal-terdistribusi	Mean, simpangan baku	Histogram, Q–Q plot
Numerik tidak normal	Lama rawat inap, biaya	Median, IQR	Histogram, boxplot
Diskrit	Jumlah kunjungan	Frekuensi, mean/median sesuai distribusi	Bar chart, histogram

Sumber tabel: Diadaptasi dari Kaliyadan dan Kulkarni (2019), Cooksey (2020), serta Ghasemi dan Zahediasl (2012).

2.15 Hubungan Jenis Variabel dengan Pemilihan Uji Statistik

Jenis variabel outcome dan desain penelitian menentukan pemilihan uji statistik inferensial (Chicco *et al*., 2025). Jika outcome numerik normal dibandingkan antara dua kelompok independen, independent t-test dapat digunakan (Chicco *et al*., 2025). Jika outcome numerik tidak normal atau ordinal dibandingkan antara dua kelompok independen, Mann–Whitney U test dapat digunakan (Chicco *et al*., 2025). Jika outcome numerik normal dibandingkan pada dua pengukuran berpasangan, paired t-test dapat digunakan (Chicco *et al*., 2025). Jika outcome ordinal atau numerik tidak normal dibandingkan pada dua pengukuran berpasangan, Wilcoxon signed-rank test dapat digunakan (Chicco *et al*., 2025).

Jika outcome numerik normal dibandingkan pada lebih dari dua kelompok independen, one-way ANOVA dapat digunakan (Chicco *et al*., 2025). Jika outcome ordinal atau numerik tidak normal dibandingkan pada lebih dari dua kelompok independen, Kruskal–Wallis test dapat digunakan (Chicco *et al*., 2025). Jika dua variabel kategorik dianalisis dalam tabel kontingensi, chi-square test atau Fisher's exact test dapat digunakan sesuai ukuran sampel dan expected count (Chicco *et al*., 2025). Jika dua variabel numerik dianalisis hubungannya, Pearson correlation dapat digunakan untuk hubungan linear dengan asumsi parametrik, sedangkan Spearman correlation dapat digunakan untuk data ordinal atau data tidak normal (Chicco *et al*., 2025).

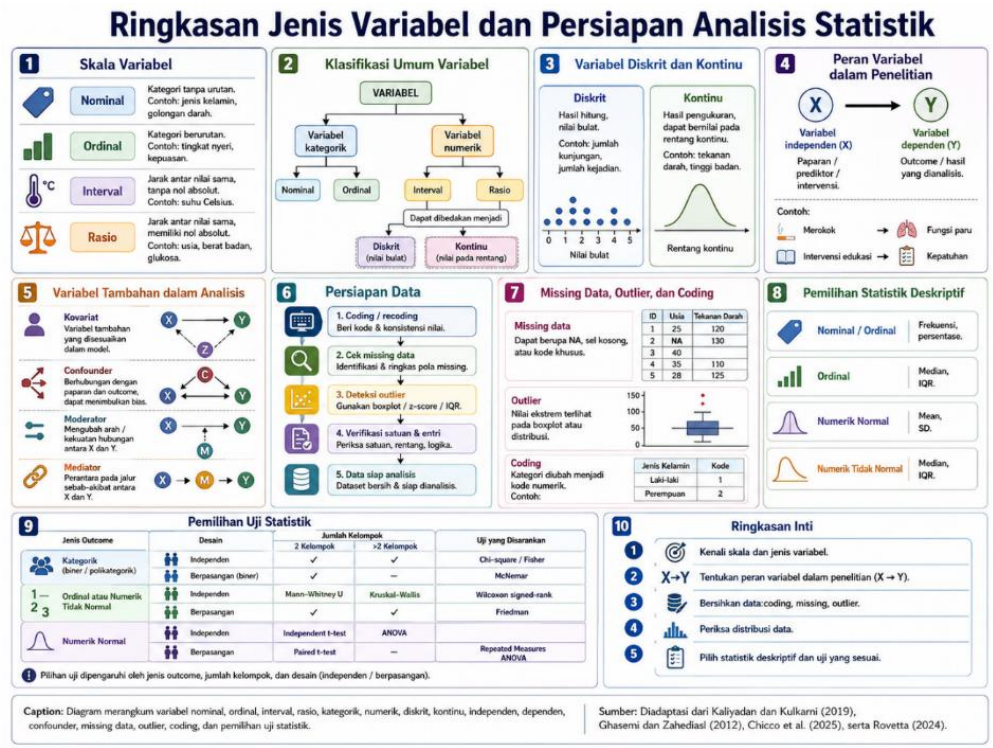


Gambar 2. 13. Pemilihan uji statistik berdasarkan variabel outcome
Diagram menunjukkan pilihan uji statistik berdasarkan outcome kategorik, ordinal, numerik normal, numerik tidak normal, jumlah kelompok, dan desain independen atau berpasangan. **Sumber:** Diadaptasi dari Chicco *et al*. (2025), Kaliyadan dan Kulkarni (2019), serta Ghasemi dan Zahediasl (2012).

Bab ini menjelaskan bahwa identifikasi jenis variabel dan skala pengukuran merupakan tahap awal yang menentukan seluruh proses analisis statistik (Kaliyadan & Kulkarni, 2019). Variabel nominal membedakan kategori tanpa urutan, sedangkan variabel ordinal membedakan kategori dengan urutan tetapi tanpa jarak yang pasti antar kategori (Kaliyadan & Kulkarni, 2019). Variabel interval memiliki jarak antar nilai yang sama tetapi tidak memiliki nol absolut, sedangkan variabel rasio memiliki jarak antar nilai yang sama dan memiliki nol absolut (Kaliyadan & Kulkarni, 2019). Variabel juga dapat diklasifikasikan menjadi kategorik, numerik, diskrit, kontinu, independen, dependen, kovariat, confounder, moderator, dan mediator sesuai peran serta karakteristik analisisnya (Kaliyadan & Kulkarni, 2019; Rovetta, 2024).

Bab ini juga menegaskan bahwa jenis variabel menentukan statistik deskriptif dan uji inferensial yang sesuai (Chicco *et al*., 2025).

Variabel kategorik umumnya disajikan dengan frekuensi dan persentase, sedangkan variabel numerik disajikan dengan mean-simpangan baku atau median-interquartile range sesuai distribusinya (Kaliyadan & Kulkarni, 2019; Ghasemi & Zahediasl, 2012). Pemeriksaan missing data, outlier, kesalahan entri, distribusi data, dan definisi operasional penting dilakukan sebelum analisis statistik inferensial (Ghasemi & Zahediasl, 2012; Rovetta, 2024). Dengan demikian, analisis statistik yang valid tidak dimulai dari memilih menu software, tetapi dari memahami variabel dan struktur data secara ilmiah (Rovetta, 2024).



Gambar 2. 14. Ringkasan jenis variabel dan persiapan analisis statistic Diagram merangkum variabel nominal, ordinal, interval, rasio, kategorik, numerik, diskrit, kontinu, independen, dependen, confounder, missing data, outlier, coding, dan pemilihan uji statistik. **Sumber:** Diadaptasi dari Kaliyadan dan Kulkarni (2019), Ghasemi dan Zahediasl (2012), Chicco *et al*. (2025), serta Rovetta (2024).

Review Questions

1. **Apa yang dimaksud variabel dalam penelitian?**
   Jawaban: Variabel adalah atribut, karakteristik, atau pengukuran yang nilainya dapat berbeda antarindividu, objek, waktu, atau kelompok penelitian.
2. **Apa perbedaan variabel nominal dan ordinal?**
   Jawaban: Variabel nominal membedakan kategori tanpa urutan, sedangkan variabel ordinal membedakan kategori yang memiliki urutan atau tingkatan.
3. **Apa contoh variabel nominal dalam penelitian kesehatan?**
   Jawaban: Contoh variabel nominal adalah jenis kelamin, golongan darah, status merokok, kelompok terapi, dan diagnosis penyakit.
4. **Apa contoh variabel ordinal?**
   Jawaban: Contoh variabel ordinal adalah tingkat nyeri, stadium penyakit, tingkat kepuasan, tingkat pendidikan, dan skor Likert.
5. **Apa perbedaan variabel interval dan rasio?**
   Jawaban: Variabel interval memiliki jarak antar nilai yang sama tetapi tidak memiliki nol absolut, sedangkan variabel rasio memiliki jarak antar nilai yang sama dan memiliki nol absolut.
6. **Apa perbedaan variabel diskrit dan kontinu?**
   Jawaban: Variabel diskrit berupa hitungan dengan nilai terpisah, sedangkan variabel kontinu dapat memiliki nilai pada rentang yang berkesinambungan.
7. **Apa perbedaan variabel independen dan dependen?**
   Jawaban: Variabel independen adalah paparan, prediktor, atau intervensi, sedangkan variabel dependen adalah outcome yang ingin dijelaskan atau dibandingkan.
8. **Mengapa data missing harus diperhatikan?**
   Jawaban: Data missing harus diperhatikan karena dapat mengurangi ukuran sampel, menurunkan kekuatan statistik, dan menimbulkan bias bila tidak terjadi secara acak.
9. **Mengapa outlier perlu diperiksa sebelum analisis?**
   Jawaban: Outlier perlu diperiksa karena dapat memengaruhi mean, simpangan baku, korelasi, regresi, dan hasil uji parametrik.
10. **Mengapa jenis variabel menentukan pemilihan uji statistik?**
    Jawaban: Jenis variabel menentukan pemilihan uji statistik karena setiap uji memiliki asumsi tertentu terkait skala data, distribusi, jumlah kelompok, dan desain penelitian.

DAFTAR PUSTAKA

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*, Article 40. https://doi.org/10.1186/s13040-025-00465-6

Cooksey, R. W. (2020). Descriptive statistics for summarising data. In *Illustrating statistical procedures: Finding meaning in quantitative data* (pp. 61–139). Springer.

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

Kaliyadan, F., & Kulkarni, V. (2019). Types of variables, descriptive statistics, and sample size. *Indian Dermatology Online Journal, 10*(1), 82–86. https://doi.org/10.4103/idoj.IDOJ_468_18

Rovetta, A. (2024). For a proper use of frequentist inferential statistics in public health and clinical research. *Global Epidemiology, 8*, 100151. https://doi.org/10.1016/j.gloepi.2024.100151

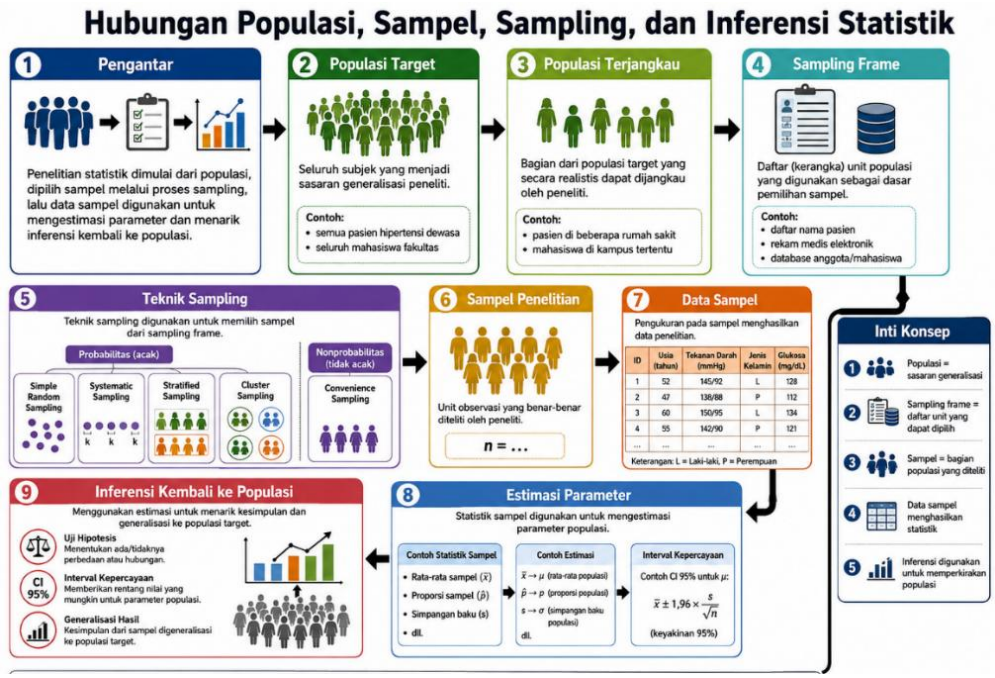
BAB 3

POPULASI, SAMPEL, TEKNIK SAMPLING, UKURAN SAMPEL, DAN KEKUATAN STATISTIK

3.1 Pendahuluan

Populasi dan sampel merupakan konsep dasar dalam statistik inferensial karena peneliti hampir selalu mengambil data dari sebagian individu untuk membuat kesimpulan mengenai kelompok yang lebih luas (Elfil & Negida, 2017). Populasi adalah keseluruhan unit yang menjadi sasaran generalisasi penelitian, sedangkan sampel adalah sebagian unit yang benar-benar diamati, diukur, atau dianalisis dalam penelitian (Elfil & Negida, 2017). Validitas kesimpulan penelitian sangat dipengaruhi oleh cara sampel dipilih dari populasi, karena metode sampling berkaitan langsung dengan validitas internal dan validitas eksternal hasil penelitian (Elfil & Negida, 2017). Ukuran sampel juga penting karena penelitian dengan sampel terlalu kecil dapat gagal mendeteksi efek yang sebenarnya ada, sedangkan sampel terlalu besar dapat membuang sumber daya dan menimbulkan persoalan etis pada penelitian klinis (Suresh & Chandrashekara, 2012; Ranganathan, 2024).

Bab ini membahas konsep populasi, target population, accessible population, sampling frame, sampel, teknik sampling probabilitas, teknik sampling nonprobabilitas, bias sampling, representativitas, ukuran sampel, power, effect size, tingkat signifikansi, kesalahan tipe I, kesalahan tipe II, dan pelaporan ukuran sampel. Dalam penelitian medis dan klinis, metode statistik perlu disesuaikan dengan tujuan penelitian, jenis data, karakteristik data, dan desain penelitian agar hasil analisis reliabel (Kosin Medical Journal, 2025). Penentuan ukuran sampel juga harus dilakukan sebelum penelitian dimulai karena perhitungan ukuran sampel merupakan bagian penting dari desain penelitian, bukan sekadar langkah administratif setelah data terkumpul (Suresh & Chandrashekara, 2012; Ranganathan, 2024).



Gambar 3. 1. Hubungan populasi, sampel, sampling, dan inferensi statistic

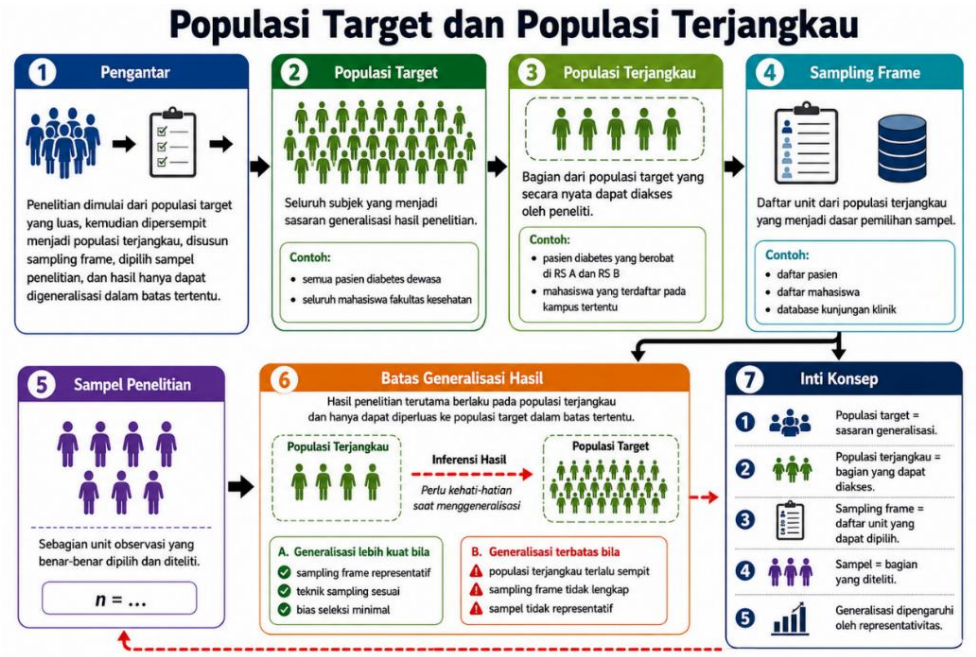
Ilustrasi menunjukkan populasi target, populasi terjangkau, sampling frame, teknik sampling, sampel penelitian, data, estimasi parameter, dan inferensi kembali ke populasi.
Sumber: Diadaptasi dari Elfil dan Negida (2017), Suresh dan Chandrashekara (2012), Ranganathan (2024), serta Kaliyadan dan Kulkarni (2019).

3.2 Populasi Target dan Populasi Terjangkau

Populasi target adalah seluruh kelompok individu, objek, atau unit yang menjadi sasaran utama kesimpulan penelitian (Elfil & Negida, 2017). Populasi target dapat berupa semua pasien diabetes tipe 2, semua ibu hamil trimester pertama, semua mahasiswa kedokteran, atau semua pasien dengan penyakit periodontal tertentu (Elfil & Negida, 2017). Populasi target biasanya lebih luas daripada kelompok yang dapat dijangkau langsung oleh peneliti (Elfil & Negida, 2017). Oleh karena itu, peneliti perlu membedakan populasi target dari populasi terjangkau agar batas generalisasi penelitian menjadi jelas (Elfil & Negida, 2017).

Populasi terjangkau adalah bagian dari populasi target yang secara praktis dapat diakses oleh peneliti berdasarkan lokasi, waktu, fasilitas, sumber daya, dan kriteria penelitian (Elfil & Negida, 2017). Misalnya, populasi target dapat berupa seluruh pasien hipertensi di Indonesia,

sedangkan populasi terjangkau dapat berupa pasien hipertensi yang datang ke tiga puskesmas tertentu selama Januari–Juni 2026 (Elfil & Negida, 2017). Perbedaan antara populasi target dan populasi terjangkau menentukan sejauh mana hasil penelitian dapat digeneralisasi (Elfil & Negida, 2017). Jika populasi terjangkau sangat berbeda dari populasi target, maka validitas eksternal penelitian dapat menurun (Elfil & Negida, 2017).



Gambar 3. 2. Populasi target dan populasi terjangkau

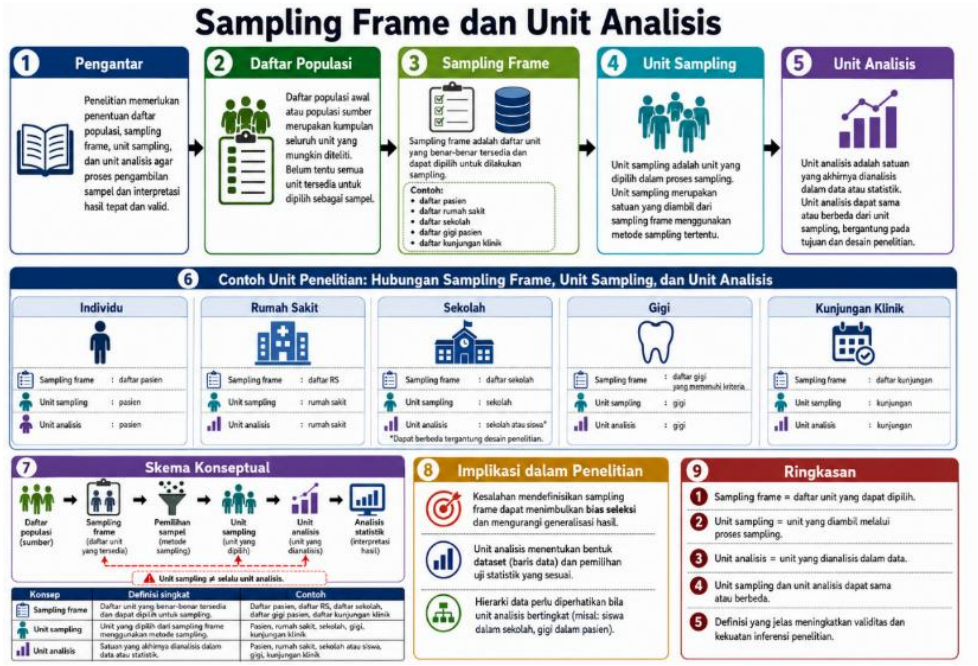
Diagram menunjukkan populasi target yang luas, populasi terjangkau yang dapat diakses peneliti, sampling frame, sampel penelitian, dan batas generalisasi hasil penelitian. **Sumber:** Diadaptasi dari Elfil dan Negida (2017) serta Spolarich (2023).

3.3 Sampling Frame dan Unit Analisis

Sampling frame adalah daftar, sistem, atau sumber yang digunakan peneliti untuk mengidentifikasi anggota populasi yang mungkin dipilih sebagai sampel (Elfil & Negida, 2017). Sampling frame dapat berupa daftar pasien, daftar mahasiswa, database rekam medis, daftar rumah tangga, registri penyakit, atau daftar fasilitas kesehatan (Elfil & Negida, 2017). Sampling frame yang tidak lengkap dapat menghasilkan coverage bias karena sebagian anggota populasi tidak memiliki peluang untuk terpilih menjadi sampel (Elfil & Negida, 2017). Oleh karena itu, kualitas

sampling frame sangat menentukan representativitas sampel penelitian (Elfil & Negida, 2017).

Unit analisis adalah satuan utama yang dianalisis dalam penelitian, seperti individu, gigi, sekolah, rumah tangga, rumah sakit, atau wilayah geografis (Kaliyadan & Kulkarni, 2019). Unit analisis harus dibedakan dari unit sampling karena unit sampling adalah unit yang dipilih dalam proses pengambilan sampel (Elfil & Negida, 2017). Dalam penelitian multilevel, unit sampling dapat berupa sekolah, sedangkan unit analisis dapat berupa siswa (Elfil & Negida, 2017). Kesalahan membedakan unit sampling dan unit analisis dapat menyebabkan kesalahan perhitungan ukuran sampel, pelanggaran independensi observasi, dan interpretasi statistik yang keliru (Kaliyadan & Kulkarni, 2019; Elfil & Negida, 2017).



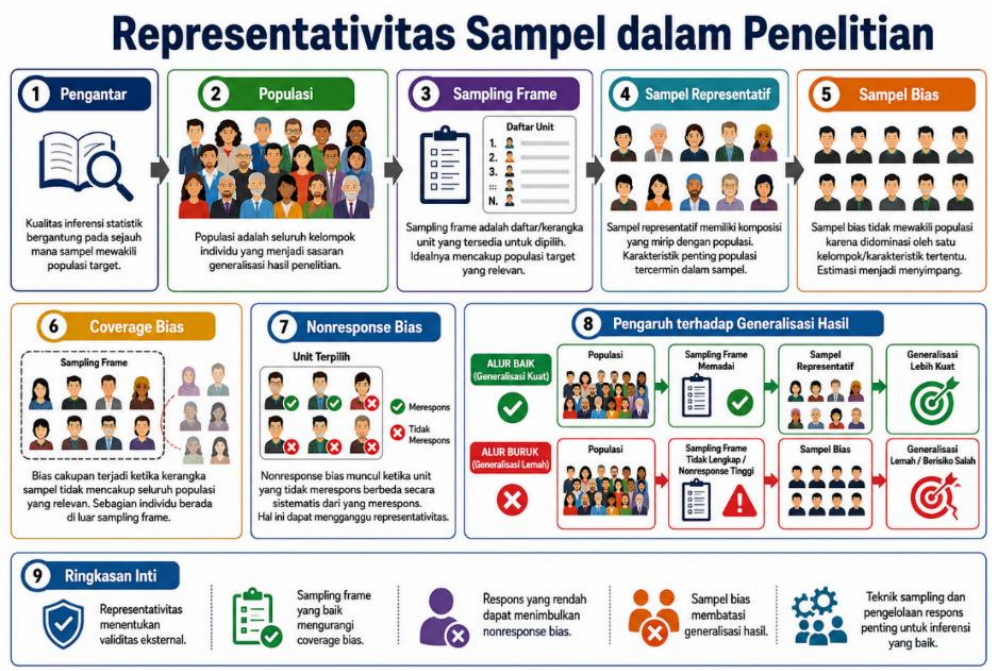
Gambar 3. 3. Sampling frame dan unit analisis

Gambar menunjukkan daftar populasi, sampling frame, unit sampling, unit analisis, serta contoh individu, rumah sakit, sekolah, gigi, dan kunjungan klinik sebagai unit penelitian. **Sumber:** Diadaptasi dari Elfil dan Negida (2017), Kaliyadan dan Kulkarni (2019), serta Spolarich (2023).

3.4 Sampel dan Representativitas

Sampel adalah bagian dari populasi yang dipilih untuk diobservasi, diukur, atau dianalisis dalam penelitian (Elfil & Negida, 2017). Sampel yang baik harus merepresentasikan populasi yang ingin digeneralisasi oleh peneliti (Elfil & Negida, 2017). Representativitas berarti karakteristik penting dalam sampel mendekati karakteristik populasi target atau populasi terjangkau (Elfil & Negida, 2017). Representativitas tidak hanya ditentukan oleh jumlah sampel, tetapi juga oleh cara pemilihan sampel, kelengkapan sampling frame, tingkat respons, dan kesesuaian kriteria inklusi-eksklusi (Elfil & Negida, 2017; Spolarich, 2023).

Sampel besar tidak selalu menjamin representativitas bila proses pemilihannya bias (Elfil & Negida, 2017). Sampel kecil dapat memberikan informasi yang berguna bila dipilih secara tepat dan sesuai dengan tujuan penelitian, tetapi ketepatan estimasinya biasanya lebih rendah dibanding sampel yang lebih besar (Suresh & Chandrashekara, 2012). Dalam statistik inferensial, representativitas sampel memengaruhi validitas eksternal, sedangkan ukuran sampel memengaruhi presisi estimasi dan kekuatan uji statistik (Suresh & Chandrashekara, 2012; Ranganathan, 2024). Oleh karena itu, peneliti harus memperhatikan kualitas sampling dan kecukupan ukuran sampel secara bersamaan (Elfil & Negida, 2017; Ranganathan, 2024).



Gambar 3. 4. Representativitas sampel dalam penelitian

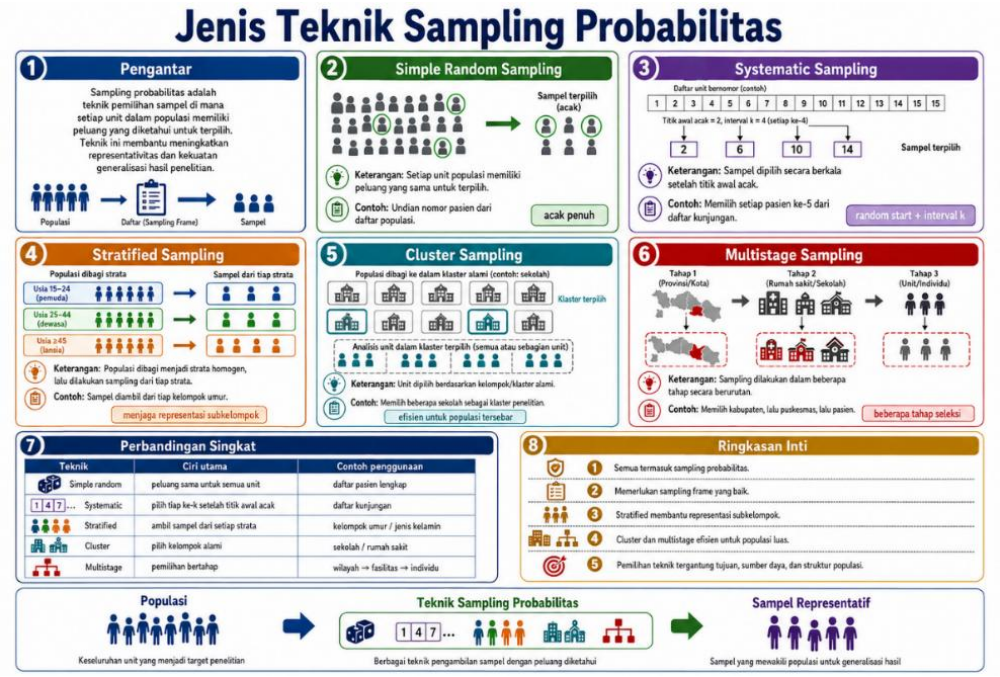
Diagram menunjukkan populasi, sampel representatif, sampel bias, sampling frame, coverage bias, nonresponse bias, dan pengaruhnya terhadap generalisasi hasil. **Sumber:** Diadaptasi dari Elfil dan Negida (2017), Spolarich (2023), serta Ranganathan (2024).

3.5 Teknik Sampling Probabilitas

Sampling probabilitas adalah teknik sampling ketika setiap anggota populasi memiliki peluang yang diketahui untuk terpilih menjadi sampel (Elfil & Negida, 2017). Teknik ini lebih mendukung generalisasi hasil karena pemilihan sampel dilakukan dengan prosedur acak atau peluang yang dapat dihitung (Elfil & Negida, 2017). Contoh sampling probabilitas adalah simple random sampling, systematic sampling, stratified sampling, cluster sampling, dan multistage sampling (Elfil & Negida, 2017). Sampling probabilitas sering dipilih ketika peneliti memiliki sampling frame yang jelas dan ingin memperkirakan parameter populasi secara lebih valid (Elfil & Negida, 2017).

Kelebihan sampling probabilitas adalah kemampuan mengurangi selection bias dan meningkatkan peluang representativitas sampel (Elfil & Negida, 2017). Keterbatasannya adalah kebutuhan sampling frame yang lengkap, biaya yang lebih besar, prosedur yang lebih kompleks, dan kemungkinan sulit dilakukan pada populasi tersembunyi atau sulit dijangkau (Elfil & Negida, 2017). Dalam penelitian klinis berbasis rumah

sakit, sampling probabilitas kadang sulit dilakukan karena pasien datang berdasarkan akses layanan, rujukan, dan kriteria klinis tertentu (Elfil & Negida, 2017). Oleh karena itu, teknik sampling harus dipilih sesuai tujuan penelitian, sumber daya, dan struktur populasi yang tersedia (Elfil & Negida, 2017).



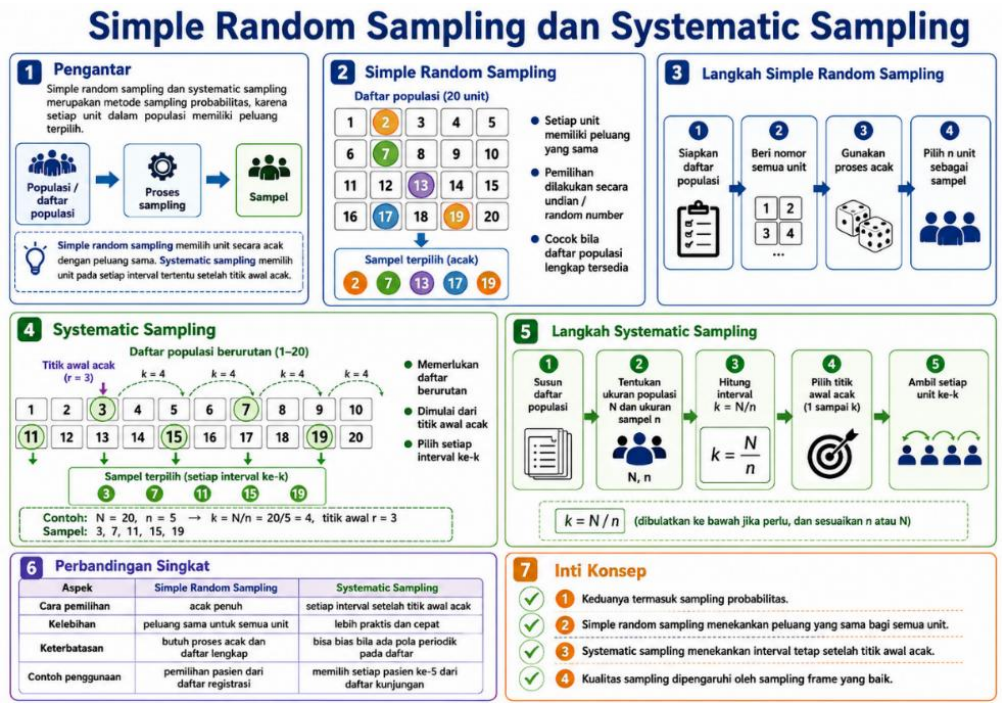
Gambar 3. 5. Jenis teknik sampling probabilitas

Gambar menunjukkan simple random sampling, systematic sampling, stratified sampling, cluster sampling, dan multistage sampling sebagai teknik sampling probabilitas. **Sumber:** Diadaptasi dari Elfil dan Negida (2017) serta Spolarich (2023).

3.6 Simple Random Sampling dan Systematic Sampling

Simple random sampling adalah teknik sampling ketika setiap anggota populasi memiliki peluang yang sama untuk terpilih sebagai sampel (Elfil & Negida, 2017). Teknik ini dapat dilakukan dengan tabel angka acak, generator angka acak, atau perangkat lunak statistik (Elfil & Negida, 2017). Simple random sampling memerlukan sampling frame yang lengkap agar semua anggota populasi memiliki peluang yang sama untuk dipilih (Elfil & Negida, 2017). Jika sampling frame tidak lengkap, simple random sampling tetap dapat menghasilkan bias karena sebagian anggota populasi tidak masuk daftar pemilihan (Elfil & Negida, 2017).

Systematic sampling adalah teknik sampling ketika peneliti memilih sampel setiap interval tertentu dari daftar populasi setelah menentukan titik awal secara acak (Elfil & Negida, 2017). Interval sampling biasanya diperoleh dengan membagi jumlah populasi dengan jumlah sampel yang dibutuhkan (Elfil & Negida, 2017). Teknik ini lebih sederhana daripada simple random sampling ketika daftar populasi sangat panjang (Elfil & Negida, 2017). Namun, systematic sampling dapat bias bila daftar populasi memiliki pola periodik yang berhubungan dengan interval pemilihan (Elfil & Negida, 2017).



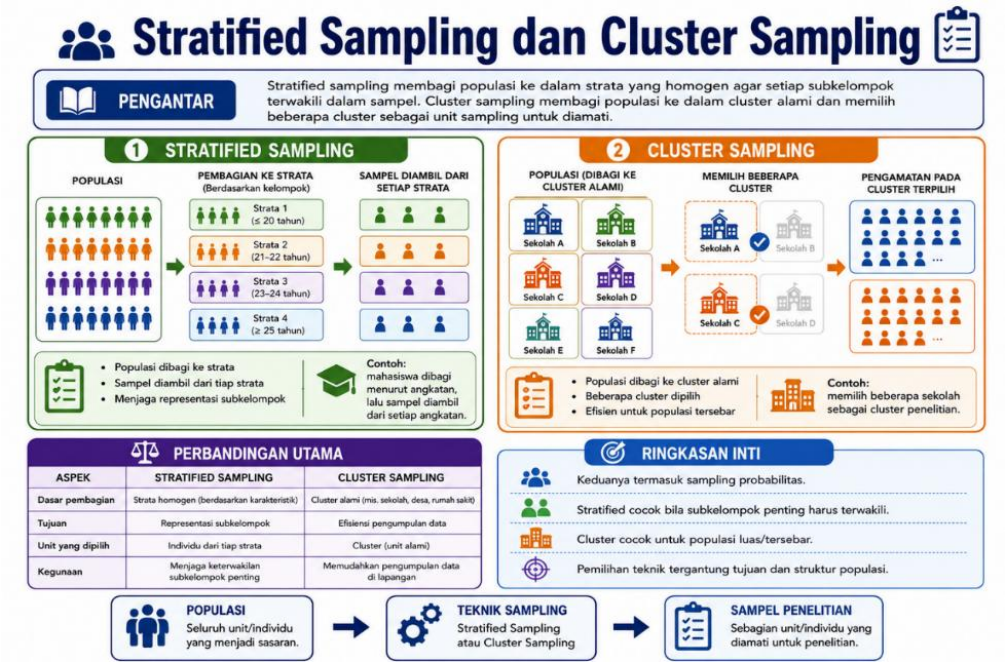
Gambar 3. 6. Simple random sampling dan systematic sampling
Diagram menunjukkan pemilihan acak dari seluruh daftar populasi dan pemilihan setiap interval tertentu setelah titik awal acak. **Sumber:** Diadaptasi dari Elfil dan Negida (2017).

3.7 Stratified Sampling dan Cluster Sampling

Stratified sampling adalah teknik sampling ketika populasi dibagi terlebih dahulu menjadi strata berdasarkan karakteristik penting, kemudian sampel dipilih dari setiap strata (Elfil & Negida, 2017). Strata dapat berupa jenis kelamin, kelompok usia, tingkat penyakit, lokasi geografis, atau kategori risiko klinis (Elfil & Negida, 2017). Teknik ini berguna ketika peneliti ingin memastikan bahwa subkelompok penting

dalam populasi terwakili dalam sampel (Elfil & Negida, 2017). Stratified sampling dapat meningkatkan presisi estimasi bila strata homogen di dalam kelompok dan berbeda antar kelompok (Elfil & Negida, 2017).

Cluster sampling adalah teknik sampling ketika populasi dibagi menjadi kelompok alami atau cluster, kemudian beberapa cluster dipilih untuk diteliti (Elfil & Negida, 2017). Cluster dapat berupa sekolah, desa, klinik, rumah sakit, atau wilayah administratif (Elfil & Negida, 2017). Cluster sampling sering digunakan ketika sampling frame individu sulit diperoleh, tetapi daftar cluster tersedia (Elfil & Negida, 2017). Kelemahan cluster sampling adalah observasi dalam cluster cenderung lebih mirip satu sama lain, sehingga analisis perlu memperhitungkan design effect atau intracluster correlation (Elfil & Negida, 2017; Suresh & Chandrashekara, 2012).



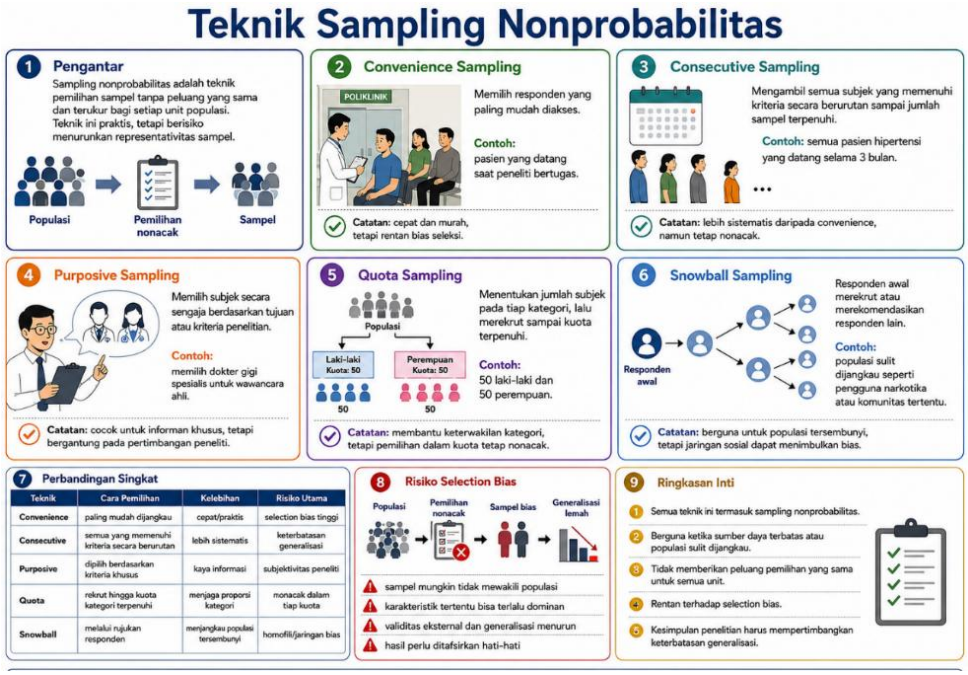
Gambar 3. 7. Stratified sampling dan cluster sampling

Gambar menunjukkan pembagian populasi ke dalam strata untuk memastikan representasi subkelompok dan pembagian populasi ke dalam cluster sebagai unit sampling. **Sumber:** Diadaptasi dari Elfil dan Negida (2017), Suresh dan Chandrashekara (2012), serta Spolarich (2023).

3.8 Teknik Sampling Nonprobabilitas

Sampling nonprobabilitas adalah teknik sampling ketika peluang setiap anggota populasi untuk terpilih tidak diketahui secara pasti (Elfil & Negida, 2017). Contoh sampling nonprobabilitas adalah convenience sampling, consecutive sampling, purposive sampling, quota sampling, dan snowball sampling (Elfil & Negida, 2017). Teknik ini sering digunakan ketika sampling frame tidak tersedia, populasi sulit dijangkau, atau penelitian bersifat eksploratif (Elfil & Negida, 2017). Sampling nonprobabilitas lebih mudah dilakukan, tetapi memiliki keterbatasan dalam generalisasi hasil karena risiko selection bias lebih tinggi (Elfil & Negida, 2017).

Convenience sampling memilih subjek yang paling mudah dijangkau oleh peneliti (Elfil & Negida, 2017). Consecutive sampling memasukkan semua subjek yang memenuhi kriteria secara berurutan selama periode tertentu sampai jumlah sampel terpenuhi (Elfil & Negida, 2017). Purposive sampling memilih subjek berdasarkan karakteristik tertentu yang relevan dengan tujuan penelitian (Elfil & Negida, 2017). Snowball sampling digunakan ketika peserta awal membantu merekrut peserta lain, terutama pada populasi tersembunyi atau sulit dijangkau (Elfil & Negida, 2017).



Gambar 3. 8. Teknik sampling nonprobabilitas

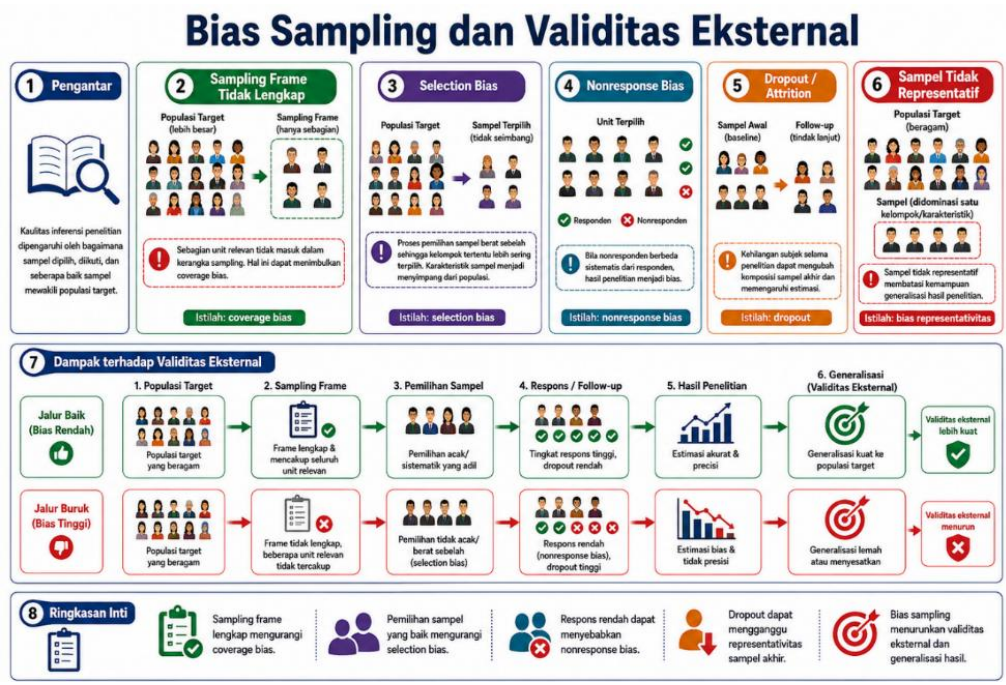
Diagram menunjukkan convenience sampling, consecutive sampling, purposive sampling, quota sampling, dan snowball sampling serta risiko selection bias. **Sumber:** Diadaptasi dari Elfil dan Negida (2017) serta Spolarich (2023).

3.9 Bias Sampling dan Validitas Eksternal

Bias sampling terjadi ketika sampel yang dipilih berbeda secara sistematis dari populasi yang ingin digeneralisasi (Elfil & Negida, 2017). Bias ini dapat muncul akibat sampling frame yang tidak lengkap, pemilihan subjek yang tidak acak, tingkat respons rendah, dropout tidak seimbang, atau kriteria inklusi yang terlalu sempit (Elfil & Negida, 2017). Selection bias dapat menyebabkan estimasi prevalensi, rata-rata, risiko, atau efek terapi menjadi tidak sesuai dengan kondisi populasi sebenarnya (Elfil & Negida, 2017). Oleh karena itu, peneliti perlu menjelaskan metode sampling dan karakteristik peserta secara transparan dalam laporan penelitian (Hopewell *et al.*, 2025).

Validitas eksternal adalah sejauh mana hasil penelitian dapat diterapkan pada populasi, setting, waktu, atau kondisi lain di luar sampel penelitian (Elfil & Negida, 2017). Validitas eksternal dipengaruhi oleh metode sampling, karakteristik sampel, kriteria inklusi-eksklusi, setting penelitian, dan tingkat respons (Elfil & Negida, 2017). Sampel yang sangat selektif dapat meningkatkan kontrol internal, tetapi dapat membatasi

generalisasi hasil ke pasien dunia nyata (Elfil & Negida, 2017). Dengan demikian, peneliti harus menyeimbangkan kebutuhan kontrol metodologis dan kebutuhan generalisasi hasil penelitian (Elfil & Negida, 2017).



Gambar 3. 9. Bias sampling dan validitas eksternal

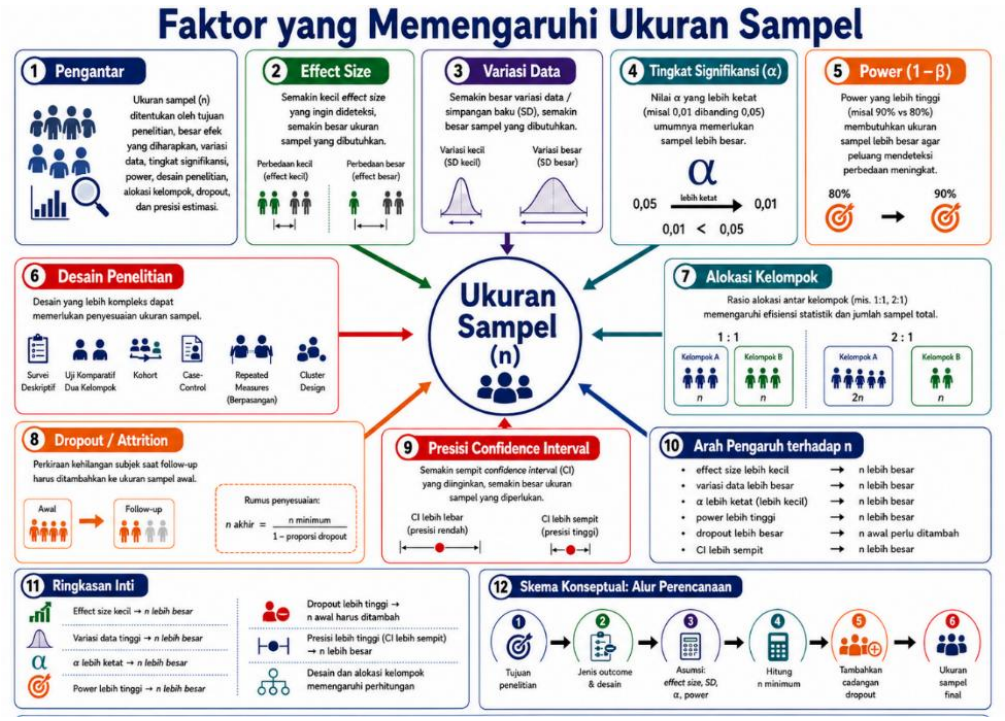
Gambar menunjukkan sampling frame tidak lengkap, selection bias, nonresponse bias, dropout, sampel tidak representatif, dan dampaknya terhadap validitas eksternal. **Sumber:** Diadaptasi dari Elfil dan Negida (2017), Hopewell *et al*. (2025), serta Spolarich (2023).

3.10 Ukuran Sampel dan Presisi Estimasi

Ukuran sampel adalah jumlah unit observasi yang dibutuhkan untuk menjawab pertanyaan penelitian dengan tingkat presisi dan kekuatan statistik yang memadai (Suresh & Chandrashekara, 2012). Perhitungan ukuran sampel bergantung pada tujuan penelitian, jenis outcome, desain penelitian, effect size yang diharapkan, variasi data, tingkat signifikansi, power, dan kemungkinan dropout (Suresh & Chandrashekara, 2012; Ranganathan, 2024). Dalam studi estimasi, ukuran sampel berhubungan erat dengan lebar confidence interval karena sampel yang lebih besar umumnya menghasilkan estimasi yang lebih presisi (Suresh & Chandrashekara, 2012). Dalam studi uji hipotesis,

ukuran sampel berhubungan erat dengan kemampuan mendeteksi efek yang bermakna secara statistik (Suresh & Chandrashekara, 2012).

Ukuran sampel yang terlalu kecil dapat menghasilkan confidence interval lebar dan meningkatkan risiko kesalahan tipe II (Suresh & Chandrashekara, 2012). Ukuran sampel yang terlalu besar dapat membuat perbedaan kecil menjadi signifikan secara statistik meskipun tidak bermakna secara klinis (Rovetta, 2024). Oleh karena itu, ukuran sampel harus ditentukan berdasarkan pertanyaan penelitian dan besar efek yang dianggap penting secara ilmiah atau klinis (Ranganathan, 2024). Perhitungan ukuran sampel juga perlu mempertimbangkan dropout, missing data, atau kehilangan follow-up agar jumlah akhir peserta tetap mencukupi untuk analisis utama (Ranganathan, 2024).



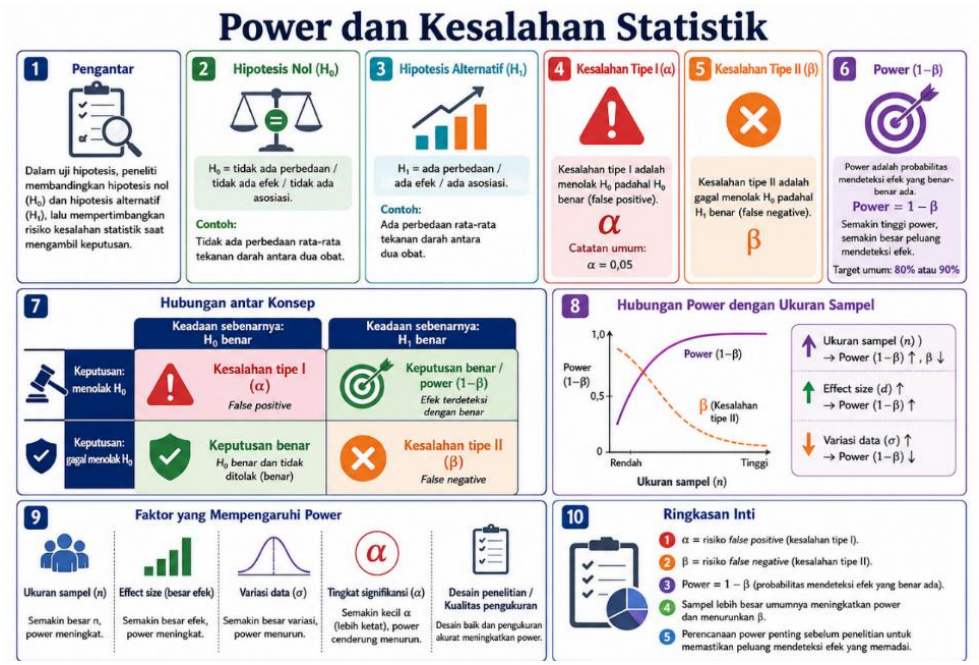
Gambar 3. 10. Faktor yang memengaruhi ukuran sampel

Diagram menunjukkan effect size, variasi data, tingkat signifikansi, power, desain penelitian, alokasi kelompok, dropout, dan presisi confidence interval sebagai faktor penentu ukuran sampel. **Sumber:** Diadaptasi dari Suresh dan Chandrashekara (2012), Ranganathan (2024), serta Bujang (2024).

3.11 Power, Kesalahan Tipe I, dan Kesalahan Tipe II

Power adalah probabilitas penelitian mendeteksi efek yang benar-benar ada dalam populasi (Suresh & Chandrashekara, 2012). Power sering dinyatakan sebagai 1−β, dengan β sebagai probabilitas kesalahan tipe II (Suresh & Chandrashekara, 2012). Kesalahan tipe I terjadi ketika peneliti menolak hipotesis nol yang sebenarnya benar (Rovetta, 2024). Kesalahan tipe II terjadi ketika peneliti gagal menolak hipotesis nol padahal efek sebenarnya ada (Suresh & Chandrashekara, 2012).

Tingkat signifikansi atau α sering ditetapkan pada 0,05 dalam banyak penelitian, tetapi nilai ini harus dipahami sebagai konvensi metodologis, bukan batas mutlak kebenaran ilmiah (Rovetta, 2024). Power yang umum digunakan dalam perencanaan penelitian adalah 80% atau 90%, tergantung konteks ilmiah, konsekuensi klinis, dan sumber daya penelitian (Suresh & Chandrashekara, 2012). Semakin tinggi power yang diinginkan, semakin besar ukuran sampel yang diperlukan (Suresh & Chandrashekara, 2012). Semakin kecil effect size yang ingin dideteksi, semakin besar ukuran sampel yang diperlukan (Suresh & Chandrashekara, 2012).



Gambar 3. 11. Power dan kesalahan statistic

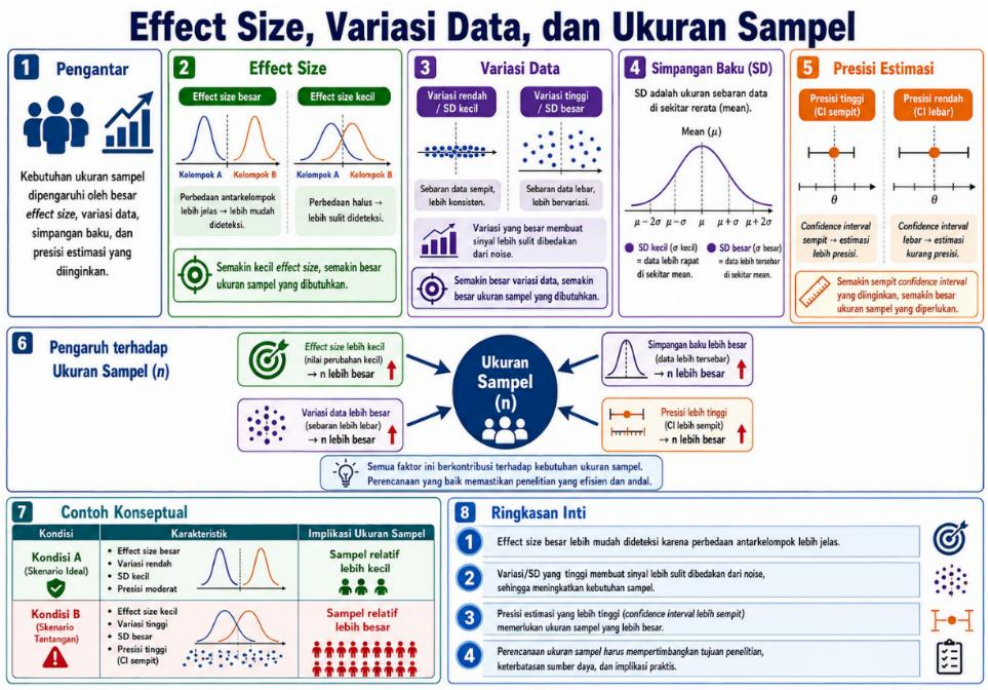
Gambar menunjukkan hipotesis nol, hipotesis alternatif, α, β, power, kesalahan tipe I, kesalahan tipe II, dan hubungan power dengan ukuran

sampel. **Sumber:** Diadaptasi dari Suresh dan Chandrashekara (2012), Rovetta (2024), serta Ranganathan (2024).

3.12 Effect Size dan Variabilitas Data

Effect size adalah besar perbedaan, hubungan, atau efek yang ingin dideteksi dalam penelitian (Suresh & Chandrashekara, 2012). Effect size dapat berupa perbedaan rata-rata, perbedaan proporsi, rasio risiko, odds ratio, hazard ratio, koefisien korelasi, atau ukuran efek standar seperti Cohen's d (Suresh & Chandrashekara, 2012). Effect size yang lebih besar membutuhkan ukuran sampel lebih kecil untuk dideteksi, sedangkan effect size yang lebih kecil membutuhkan ukuran sampel lebih besar (Suresh & Chandrashekara, 2012). Oleh karena itu, effect size harus ditentukan berdasarkan literatur, studi pendahuluan, kepentingan klinis, atau pertimbangan ilmiah yang jelas (Ranganathan, 2024).

Variabilitas data juga memengaruhi ukuran sampel karena data yang lebih bervariasi membutuhkan sampel lebih besar untuk menghasilkan estimasi yang presisi (Suresh & Chandrashekara, 2012). Simpangan baku yang lebih besar pada outcome numerik akan meningkatkan ukuran sampel yang dibutuhkan untuk mendeteksi perbedaan rata-rata tertentu (Suresh & Chandrashekara, 2012). Pada outcome proporsi, ukuran sampel dipengaruhi oleh proporsi yang diharapkan dan presisi yang diinginkan (Suresh & Chandrashekara, 2012). Dengan demikian, informasi dari studi sebelumnya sangat penting untuk memperkirakan variabilitas dan effect size secara realistis (Ranganathan, 2024).



Gambar 3. 12. Effect size, variasi data, dan ukuran sampel

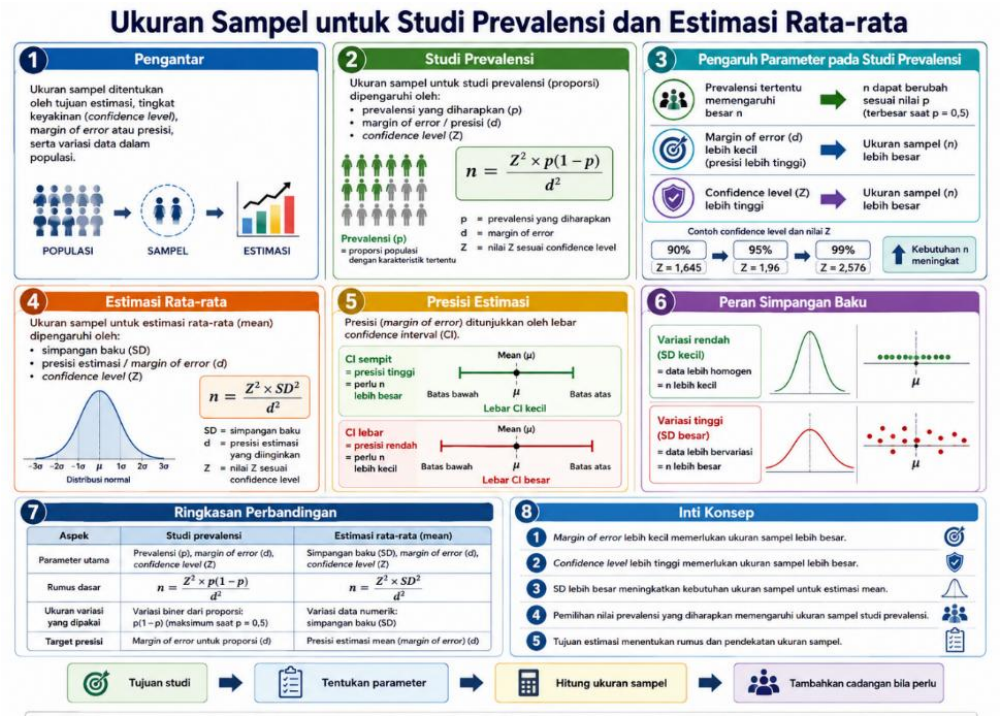
Diagram menunjukkan effect size besar versus kecil, variasi rendah versus tinggi, simpangan baku, presisi estimasi, dan konsekuensinya terhadap kebutuhan ukuran sampel. **Sumber:** Diadaptasi dari Suresh dan Chandrashekara (2012), Ranganathan (2024), serta Bujang (2024).

3.13 Ukuran Sampel untuk Studi Deskriptif dan Estimasi Prevalensi

Studi deskriptif sering bertujuan memperkirakan nilai rata-rata, proporsi, atau prevalensi dalam populasi tertentu (Suresh & Chandrashekara, 2012). Ukuran sampel untuk estimasi prevalensi dipengaruhi oleh prevalensi yang diharapkan, tingkat kepercayaan, dan presisi atau margin of error yang diinginkan (Suresh & Chandrashekara, 2012). Jika prevalensi awal belum diketahui, peneliti dapat menggunakan estimasi konservatif atau merujuk pada studi pendahuluan yang paling relevan (Suresh & Chandrashekara, 2012). Presisi yang lebih ketat membutuhkan ukuran sampel lebih besar (Suresh & Chandrashekara, 2012).

Dalam studi yang mengestimasi rata-rata, ukuran sampel dipengaruhi oleh simpangan baku populasi yang diharapkan, tingkat kepercayaan, dan margin of error yang dapat diterima (Suresh & Chandrashekara, 2012). Jika peneliti ingin confidence interval yang sempit, ukuran sampel harus ditingkatkan (Suresh & Chandrashekara,

2012). Perhitungan ukuran sampel dalam studi deskriptif sebaiknya dikaitkan dengan tujuan estimasi, bukan hanya angka minimal yang umum digunakan (Ranganathan, 2024). Oleh karena itu, peneliti perlu menyatakan parameter yang ingin diestimasi, asumsi yang digunakan, dan margin of error yang dianggap dapat diterima (Ranganathan, 2024).



Gambar 3. 13. Ukuran sampel untuk studi prevalensi dan estimasi rata-rata

Gambar menunjukkan prevalensi yang diharapkan, margin of error, confidence level, simpangan baku, presisi estimasi, dan kebutuhan ukuran sampel.

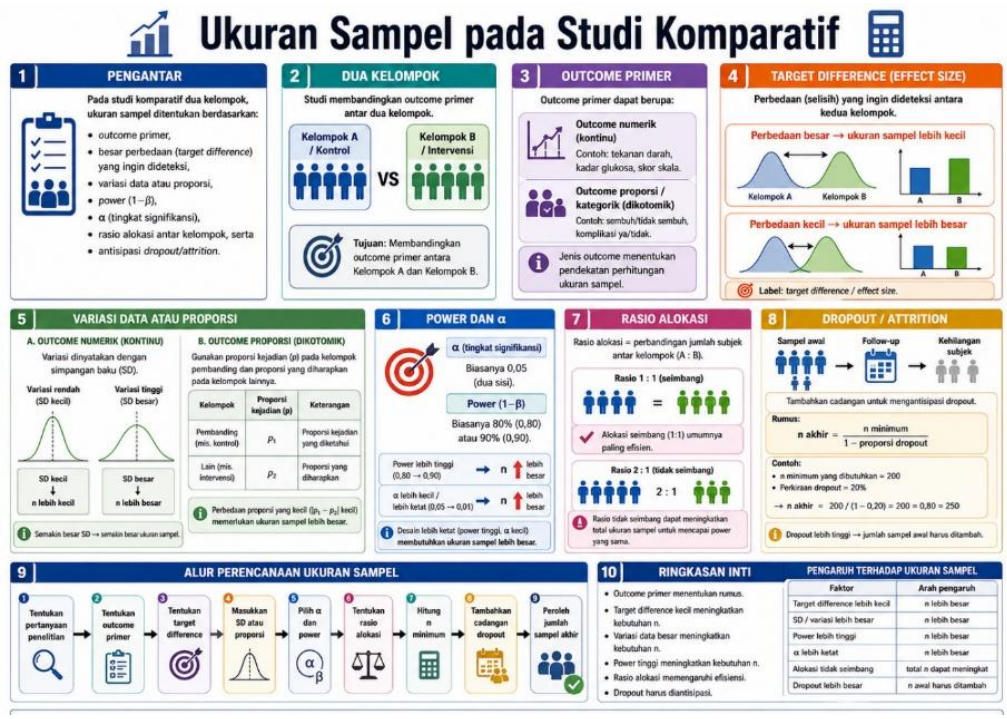
Sumber: Diadaptasi dari Suresh dan Chandrashekara (2012) serta Ranganathan (2024).

3.14 Ukuran Sampel untuk Studi Komparatif

Studi komparatif bertujuan membandingkan outcome antara dua kelompok atau lebih (Chicco *et al.*, 2025). Ukuran sampel untuk perbandingan rata-rata dipengaruhi oleh perbedaan rata-rata yang ingin dideteksi, simpangan baku, tingkat signifikansi, power, dan rasio alokasi kelompok (Suresh & Chandrashekara, 2012). Ukuran sampel untuk perbandingan proporsi dipengaruhi oleh proporsi kejadian pada tiap kelompok, perbedaan proporsi yang dianggap bermakna, tingkat

signifikansi, power, dan rasio alokasi kelompok (Suresh & Chandrashekara, 2012). Studi dengan perbedaan antar kelompok yang kecil membutuhkan ukuran sampel lebih besar dibanding studi dengan perbedaan antar kelompok yang besar (Suresh & Chandrashekara, 2012).

Dalam uji klinis acak, pelaporan ukuran sampel harus menjelaskan bagaimana ukuran sampel ditentukan dan asumsi apa saja yang mendukung perhitungan tersebut (Hopewell *et al*., 2025). CONSORT 2025 menekankan bahwa laporan randomized trial perlu menyertakan informasi penting agar pembaca dapat menilai, menginterpretasi, dan menggunakan hasil penelitian secara kritis (Hopewell *et al*., 2025). Peneliti perlu melaporkan outcome primer, target difference, tingkat signifikansi, power, asumsi variabilitas atau proporsi, rasio alokasi, dan penyesuaian untuk dropout bila digunakan (Hopewell *et al*., 2025; Ranganathan, 2024). Dengan demikian, ukuran sampel bukan hanya angka, tetapi pernyataan metodologis yang harus transparan dan dapat diaudit (Hopewell *et al*., 2025).



Gambar 3. 14. Ukuran sampel pada studi komparatif

Diagram menunjukkan dua kelompok, outcome primer, target difference, simpangan baku atau proporsi, power, α, rasio alokasi, dropout, dan jumlah sampel akhir. **Sumber:** Diadaptasi dari Suresh dan

Chandrashekara (2012), Ranganathan (2024), Chicco *et al*. (2025), serta Hopewell *et al*. (2025).

3.15 Ukuran Sampel untuk Studi Korelasi dan Regresi

Studi korelasi bertujuan menilai kekuatan dan arah hubungan antara dua variabel (Chicco *et al*., 2025). Ukuran sampel untuk korelasi dipengaruhi oleh koefisien korelasi yang diharapkan, tingkat signifikansi, power, dan presisi confidence interval (Bujang, 2024). Korelasi yang lebih kecil membutuhkan ukuran sampel lebih besar agar dapat dideteksi secara statistik (Bujang, 2024). Bujang (2024) membahas perhitungan ukuran sampel untuk Pearson correlation, Spearman rank correlation, dan Kendall's Tau-b berdasarkan effect size serta lebar confidence interval (Bujang, 2024).

Studi regresi membutuhkan perhatian khusus karena jumlah prediktor, distribusi outcome, multikolinearitas, event per variable, dan kompleksitas model dapat memengaruhi kebutuhan ukuran sampel (Suresh & Chandrashekara, 2012). Ukuran sampel yang terlalu kecil pada model regresi dapat menghasilkan estimasi yang tidak stabil dan confidence interval yang lebar (Suresh & Chandrashekara, 2012). Dalam regresi logistik, jumlah kejadian outcome sering menjadi pertimbangan penting karena model dengan terlalu sedikit kejadian dapat mengalami overfitting (Suresh & Chandrashekara, 2012). Oleh karena itu, rencana analisis regresi sebaiknya ditentukan sebelum penelitian dan disertai justifikasi ukuran sampel yang sesuai (Ranganathan, 2024).



Gambar 3. 15. Ukuran sampel untuk korelasi dan regresi

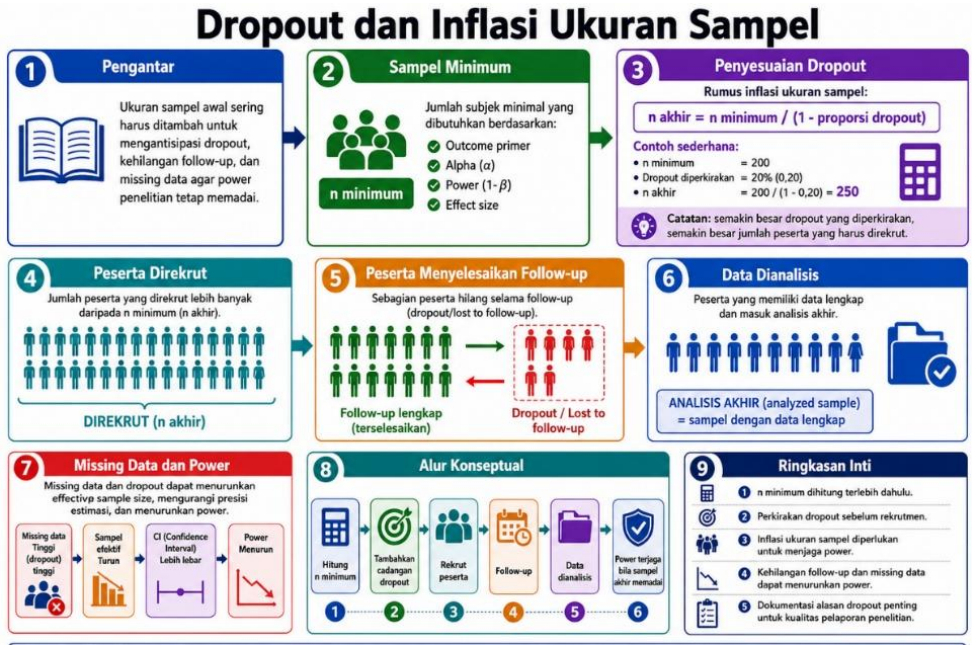
Gambar menunjukkan koefisien korelasi, confidence interval, jumlah prediktor, event per variable, risiko overfitting, dan kebutuhan ukuran sampel pada model korelasi-regresi. **Sumber:** Diadaptasi dari Bujang (2024), Suresh dan Chandrashekara (2012), serta Ranganathan (2024).

3.16 Dropout, Missing Data, dan Inflasi Ukuran Sampel

Dropout adalah kondisi ketika peserta yang telah direkrut tidak menyelesaikan penelitian atau tidak memiliki data outcome akhir (Ranganathan, 2024). Missing data dapat terjadi akibat peserta tidak hadir, menolak pemeriksaan, alat ukur gagal, catatan hilang, atau follow-up tidak lengkap (Ranganathan, 2024). Dropout dan missing data dapat menurunkan power serta meningkatkan risiko bias bila terjadi secara tidak acak (Ranganathan, 2024). Oleh karena itu, perhitungan ukuran sampel sering perlu ditambah untuk mengantisipasi dropout atau kehilangan follow-up (Ranganathan, 2024).

Jika peneliti memperkirakan dropout sebesar 10%, maka jumlah peserta yang direkrut biasanya perlu lebih besar daripada jumlah sampel minimum untuk analisis utama (Ranganathan, 2024). Penyesuaian dropout harus didasarkan pada pengalaman studi sebelumnya, karakteristik populasi, durasi follow-up, dan beban partisipasi peserta (Ranganathan, 2024). CONSORT 2025 juga menekankan pentingnya diagram alur peserta dalam randomized trial agar pembaca dapat melihat jumlah peserta yang dinilai, dirandomisasi, menerima intervensi,

menyelesaikan follow-up, dan dianalisis (Hopewell *et al*., 2025). Dengan demikian, transparansi alur peserta merupakan bagian penting dari pelaporan kualitas penelitian (Hopewell *et al*., 2025).



Gambar 3. 16. Dropout dan inflasi ukuran sampel

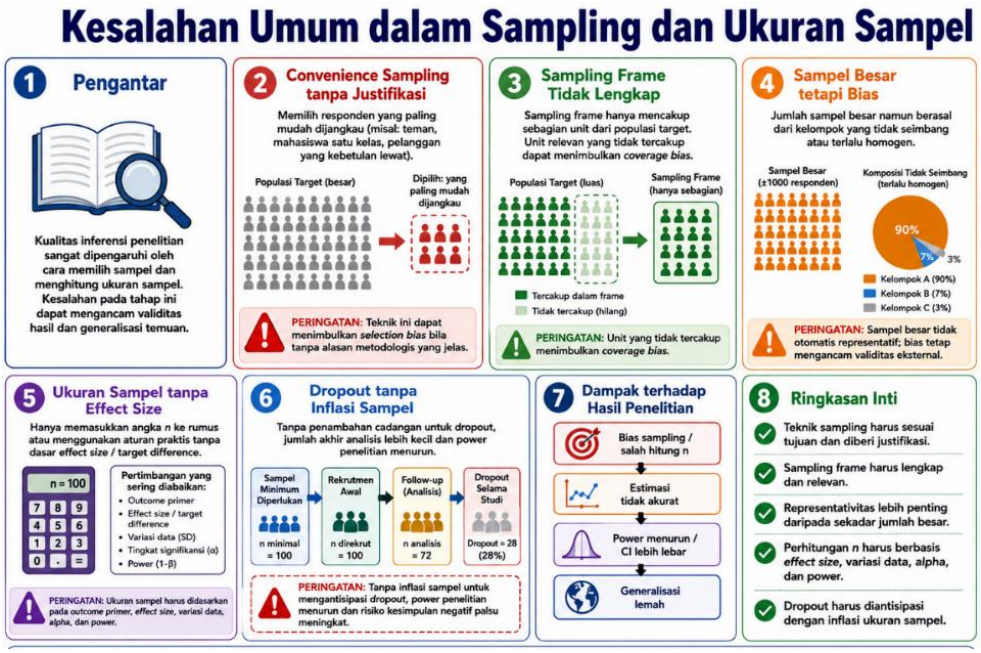
Diagram menunjukkan sampel minimum, penyesuaian dropout, peserta direkrut, peserta menyelesaikan follow-up, data dianalisis, dan pengaruh missing data terhadap power.
Sumber: Diadaptasi dari Ranganathan (2024), Suresh dan Chandrashekara (2012), serta Hopewell *et al*. (2025).

3.17 Kesalahan Umum dalam Sampling dan Ukuran Sampel

Kesalahan umum pertama adalah menggunakan convenience sampling tanpa menjelaskan keterbatasan generalisasi hasil penelitian (Elfil & Negida, 2017). Kesalahan umum kedua adalah menganggap sampel besar selalu representatif meskipun sampling frame tidak lengkap atau proses pemilihan subjek bias (Elfil & Negida, 2017). Kesalahan umum ketiga adalah tidak membedakan populasi target, populasi terjangkau, sampling frame, dan sampel penelitian (Elfil & Negida, 2017). Kesalahan umum keempat adalah tidak melaporkan tingkat respons, dropout, atau alasan peserta tidak dianalisis (Hopewell *et al*., 2025).

Kesalahan umum dalam ukuran sampel adalah menggunakan aturan praktis tanpa justifikasi ilmiah (Ranganathan, 2024). Kesalahan lain adalah menghitung ukuran sampel tanpa menentukan outcome

primer dan effect size yang bermakna (Ranganathan, 2024). Peneliti juga sering tidak menambahkan inflasi untuk dropout atau kehilangan follow-up, sehingga jumlah sampel akhir menjadi kurang dari kebutuhan analisis utama (Ranganathan, 2024). Oleh karena itu, metode sampling dan ukuran sampel harus direncanakan, dijustifikasi, dan dilaporkan secara eksplisit dalam protokol maupun artikel penelitian (Hopewell *et al*., 2025; Ranganathan, 2024).

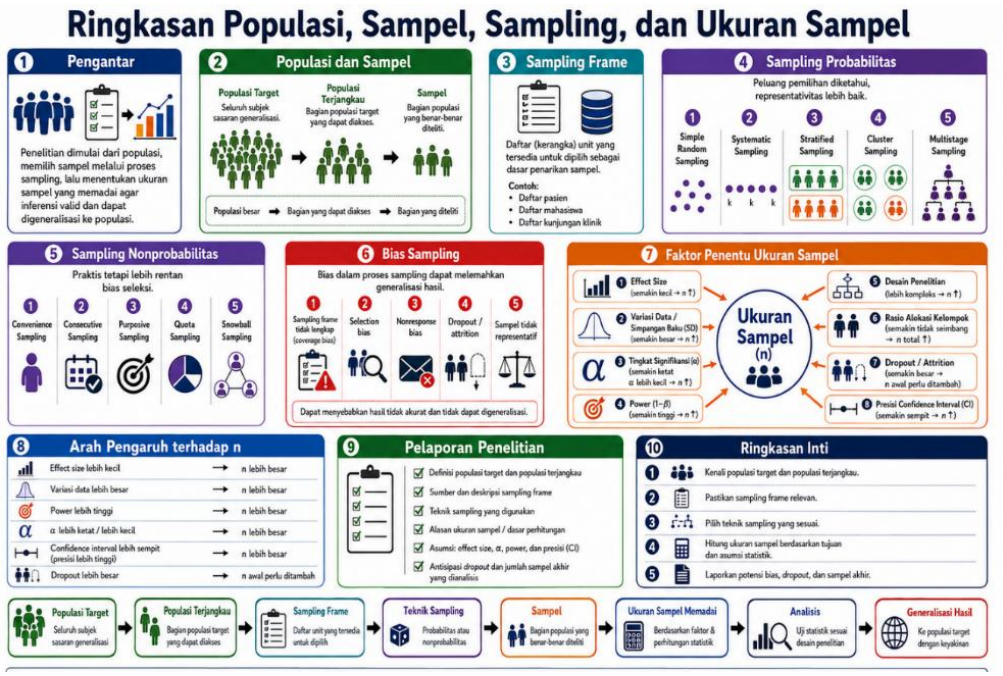


Gambar 3. 17. Kesalahan umum dalam sampling dan ukuran sampel
Gambar menunjukkan convenience sampling tanpa justifikasi, sampling frame tidak lengkap, sampel besar tetapi bias, ukuran sampel tanpa effect size, dan dropout tanpa inflasi sampel. **Sumber:** Diadaptasi dari Elfil dan Negida (2017), Ranganathan (2024), serta Hopewell *et al*. (2025).

Bab ini menjelaskan bahwa populasi, sampel, sampling frame, teknik sampling, dan ukuran sampel merupakan fondasi statistik inferensial (Elfil & Negida, 2017). Populasi target adalah kelompok luas yang menjadi sasaran generalisasi, sedangkan populasi terjangkau adalah bagian populasi yang dapat diakses secara praktis oleh peneliti (Elfil & Negida, 2017). Sampling probabilitas memberi peluang terpilih yang diketahui bagi anggota populasi, sedangkan sampling nonprobabilitas lebih mudah dilakukan tetapi memiliki risiko bias generalisasi yang lebih besar (Elfil & Negida, 2017). Representativitas sampel dipengaruhi oleh

metode sampling, kelengkapan sampling frame, tingkat respons, dropout, dan kesesuaian kriteria penelitian (Elfil & Negida, 2017; Spolarich, 2023).

Ukuran sampel harus ditentukan berdasarkan tujuan penelitian, jenis outcome, desain studi, effect size, variabilitas data, tingkat signifikansi, power, dan perkiraan dropout (Suresh & Chandrashekara, 2012; Ranganathan, 2024). Power menunjukkan kemampuan penelitian untuk mendeteksi efek yang benar-benar ada, sedangkan kesalahan tipe I dan tipe II menunjukkan risiko keputusan inferensial yang keliru (Suresh & Chandrashekara, 2012; Rovetta, 2024). Pelaporan ukuran sampel perlu menyertakan asumsi yang digunakan agar pembaca dapat menilai kualitas metodologi penelitian (Hopewell *et al*., 2025). Dengan demikian, analisis statistik yang valid tidak hanya membutuhkan uji statistik yang tepat, tetapi juga sampel yang dipilih secara benar dan ukuran sampel yang memadai (Elfil & Negida, 2017; Ranganathan, 2024).



Gambar 3. 18. Ringkasan populasi, sampel, sampling, dan ukuran sampel

Diagram merangkum populasi target, populasi terjangkau, sampling frame, sampling probabilitas, sampling nonprobabilitas, bias sampling, ukuran sampel, effect size, power, dropout, dan pelaporan penelitian.

Sumber: Diadaptasi dari Elfil dan Negida (2017), Suresh dan Chandrashekara (2012), Ranganathan (2024), Hopewell *et al*. (2025), serta Chicco *et al*. (2025).

Review Questions

1. **Apa perbedaan populasi target dan populasi terjangkau?**
   Jawaban: Populasi target adalah kelompok luas yang menjadi sasaran generalisasi, sedangkan populasi terjangkau adalah bagian dari populasi target yang dapat diakses secara praktis oleh peneliti.
2. **Apa yang dimaksud sampling frame?**
   Jawaban: Sampling frame adalah daftar, sistem, atau sumber yang digunakan untuk mengidentifikasi anggota populasi yang mungkin dipilih menjadi sampel.
3. **Apa perbedaan sampling probabilitas dan nonprobabilitas?**
   Jawaban: Sampling probabilitas memberi peluang terpilih yang diketahui bagi setiap anggota populasi, sedangkan sampling nonprobabilitas tidak memiliki peluang terpilih yang diketahui secara pasti.
4. **Apa contoh teknik sampling probabilitas?**
   Jawaban: Contohnya adalah simple random sampling, systematic sampling, stratified sampling, cluster sampling, dan multistage sampling.
5. **Apa contoh teknik sampling nonprobabilitas?**
   Jawaban: Contohnya adalah convenience sampling, consecutive sampling, purposive sampling, quota sampling, dan snowball sampling.
6. **Mengapa representativitas sampel penting?**
   Jawaban: Representativitas penting karena menentukan sejauh mana hasil penelitian dapat digeneralisasi ke populasi target.
7. **Apa yang dimaksud power dalam penelitian?**
   Jawaban: Power adalah probabilitas penelitian untuk mendeteksi efek yang benar-benar ada dalam populasi.
8. **Apa perbedaan kesalahan tipe I dan tipe II?**
   Jawaban: Kesalahan tipe I terjadi ketika peneliti menolak hipotesis nol yang sebenarnya benar, sedangkan kesalahan tipe II terjadi ketika peneliti gagal menolak hipotesis nol padahal efek sebenarnya ada.
9. **Apa faktor utama yang memengaruhi ukuran sampel?**
   Jawaban: Faktor utamanya adalah tujuan penelitian, jenis outcome, effect size, variasi data, tingkat signifikansi, power, desain penelitian, dan perkiraan dropout.

10. **Mengapa dropout perlu diperhitungkan dalam ukuran sampel?**
    Jawaban: Dropout perlu diperhitungkan karena peserta yang tidak menyelesaikan penelitian dapat menurunkan power dan membuat jumlah sampel akhir tidak mencukupi untuk analisis utama.

DAFTAR PUSTAKA

Bujang, M. A. (2024). An elaboration on sample size determination for correlation analysis: A brief guide for researchers. *Restorative Dentistry & Endodontics, 49*, e21. https://doi.org/10.5395/rde.2024.49.e21

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*, Article 40. https://doi.org/10.1186/s13040-025-00465-6

Elfil, M., & Negida, A. (2017). Sampling methods in clinical research: An educational review. *Emergency, 5*(1), e52.

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

Hopewell, S., Chan, A.-W., Collins, G. S., Hróbjartsson, A., Moher, D., Schulz, K. F., *et al*. (2025). CONSORT 2025 statement: Updated guideline for reporting randomised trials. *The Lancet*.

Kaliyadan, F., & Kulkarni, V. (2019). Types of variables, descriptive statistics, and sample size. *Indian Dermatology Online Journal, 10*(1), 82–86. https://doi.org/10.4103/idoj.IDOJ_468_18

Kosin Medical Journal. (2025). Common statistical methods used in medical research. *Kosin Medical Journal*.

Ranganathan, P. (2024). Sample size calculation in clinical research. *Perspectives in Clinical Research*.

Rovetta, A. (2024). For a proper use of frequentist inferential statistics in public health and clinical research. *Global Epidemiology, 8*, 100151. https://doi.org/10.1016/j.gloepi.2024.100151

Spolarich, A. E. (2023). Sampling methods: A guide for researchers. *Journal of Dental Hygiene, 97*(4), 73–76.

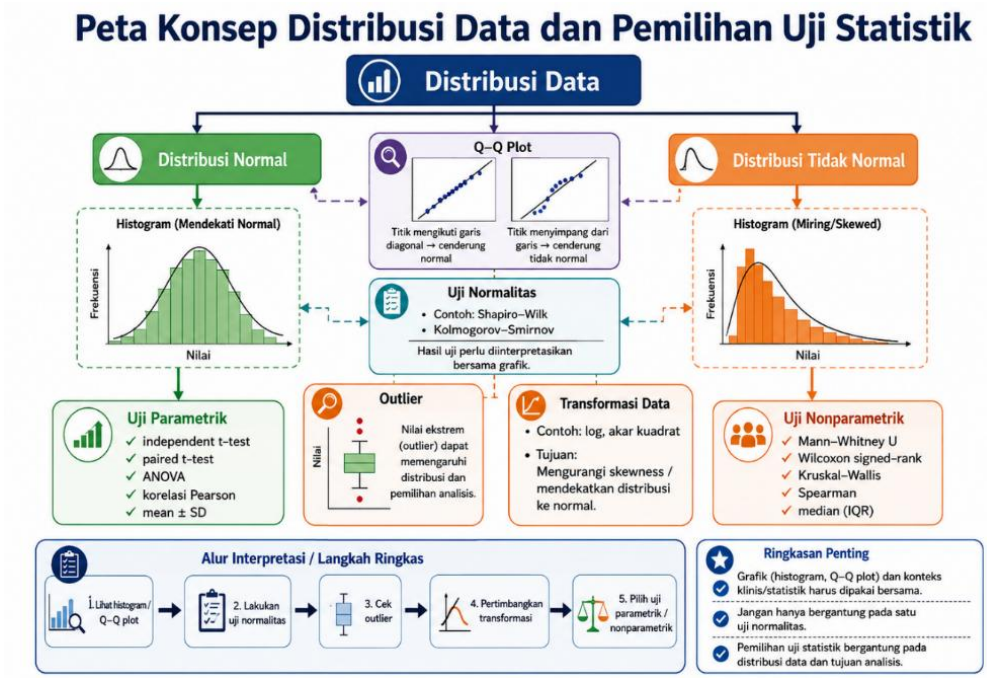
BAB 4

DISTRIBUSI DATA, UJI NORMALITAS, TRANSFORMASI DATA, DAN PEMILIHAN STATISTIK PARAMETRIK–NONPARAMETRIK

4.1 Pendahuluan

Distribusi data merupakan pola penyebaran nilai observasi dalam suatu variabel penelitian (Ghasemi & Zahediasl, 2012). Pemahaman distribusi data penting karena banyak metode statistik, seperti uji t, ANOVA, korelasi Pearson, dan regresi linear, memiliki asumsi tertentu mengenai normalitas, varians, dan bentuk hubungan antarvariabel (Ghasemi & Zahediasl, 2012; Kamath *et al*., 2025). Distribusi data dapat berbentuk normal, menceng ke kanan, menceng ke kiri, bimodal, uniform, atau memiliki outlier ekstrem (Ghasemi & Zahediasl, 2012). Pemeriksaan distribusi data membantu peneliti memilih ringkasan deskriptif yang tepat, menentukan uji parametrik atau nonparametrik, dan menghindari interpretasi statistik yang keliru (Chicco *et al*., 2025).

Bab ini membahas distribusi normal, distribusi tidak normal, skewness, kurtosis, outlier, histogram, Q–Q plot, uji Shapiro–Wilk, uji Kolmogorov–Smirnov, central limit theorem, transformasi data, statistik robust, serta logika pemilihan uji parametrik dan nonparametrik. Normalitas tidak boleh dipahami hanya sebagai hasil “signifikan” atau “tidak signifikan” dari uji normalitas, karena interpretasi distribusi data harus menggabungkan grafik, ukuran sampel, konteks klinis, dan tujuan analisis (Ghasemi & Zahediasl, 2012). Peneliti perlu menghindari keputusan mekanis berdasarkan nilai p semata karena nilai p dan confidence interval harus ditafsirkan sebagai ukuran kompatibilitas data dengan model statistik dan asumsi yang digunakan (Rovetta *et al*., 2025).



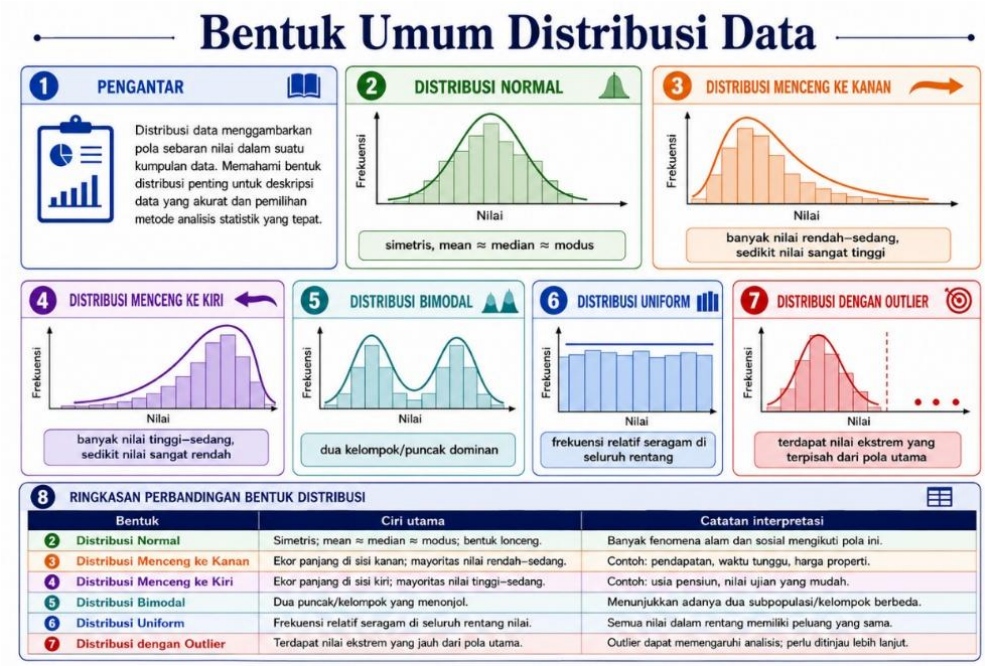
Gambar 4. 1. Peta konsep distribusi data dan pemilihan uji statistic
Ilustrasi menunjukkan hubungan antara distribusi normal, distribusi tidak normal, histogram, Q–Q plot, uji normalitas, outlier, transformasi data, uji parametrik, dan uji nonparametrik.
Sumber: Diadaptasi dari Ghasemi dan Zahediasl (2012), Chicco *et al*. (2025), Kamath *et al*. (2025), serta Rovetta *et al*. (2025).

4.2 Konsep Distribusi Data

Distribusi data adalah gambaran tentang bagaimana nilai-nilai suatu variabel tersebar dalam dataset (Cooksey, 2020). Distribusi data dapat dilihat melalui tabel frekuensi, histogram, density plot, boxplot, atau Q–Q plot (Ghasemi & Zahediasl, 2012). Distribusi yang simetris menunjukkan bahwa nilai di sekitar pusat data tersebar relatif seimbang di sisi kiri dan kanan (Ghasemi & Zahediasl, 2012). Distribusi yang tidak simetris menunjukkan adanya kemencengan atau skewness, baik ke arah kanan maupun kiri (Ghasemi & Zahediasl, 2012).

Distribusi data menentukan cara penyajian statistik deskriptif (Kaliyadan & Kulkarni, 2019). Data numerik yang relatif normal dapat diringkas dengan mean dan simpangan baku (Ghasemi & Zahediasl, 2012). Data numerik yang sangat menceng lebih tepat diringkas dengan median dan interquartile range (Cooksey, 2020). Data dengan outlier ekstrem perlu diperiksa sebelum analisis karena outlier dapat

memengaruhi mean, simpangan baku, korelasi, regresi, dan uji parametrik (Ghasemi & Zahediasl, 2012).



Gambar 4. 2. Bentuk umum distribusi data

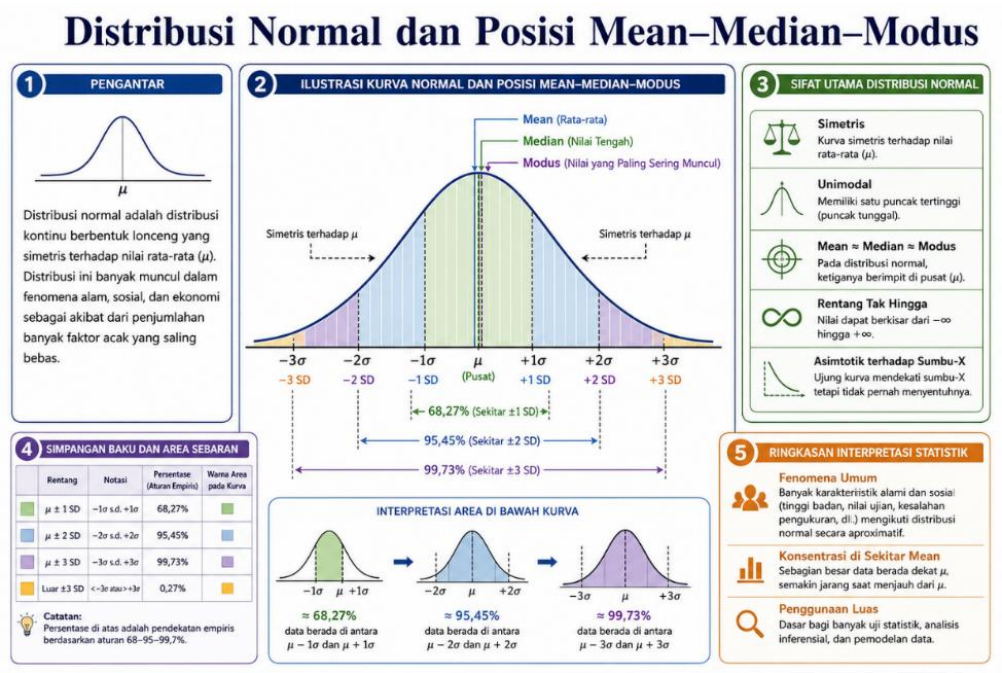
Diagram menunjukkan distribusi normal, distribusi menceng ke kanan, distribusi menceng ke kiri, distribusi bimodal, distribusi uniform, dan distribusi dengan outlier. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012), Cooksey (2020), serta Kaliyadan dan Kulkarni (2019).

4.3 Distribusi Normal

Distribusi normal adalah distribusi berbentuk lonceng yang simetris di sekitar rata-rata (Ghasemi & Zahediasl, 2012). Pada distribusi normal ideal, mean, median, dan modus berada pada posisi yang sama atau sangat berdekatan (Ghasemi & Zahediasl, 2012). Distribusi normal memiliki peran penting dalam statistik inferensial karena banyak prosedur parametrik dikembangkan berdasarkan asumsi normalitas atau distribusi sampling yang mendekati normal (Ghasemi & Zahediasl, 2012). Uji t, ANOVA, korelasi Pearson, dan regresi linear sering mengandalkan asumsi normalitas dalam bentuk tertentu, terutama pada ukuran sampel kecil atau pada residual model (Kamath *et al*., 2025).

Distribusi normal tidak berarti semua data biologis atau klinis harus sempurna berbentuk lonceng (Ghasemi & Zahediasl, 2012). Dalam praktik penelitian, data dapat mendekati normal walaupun terdapat

sedikit penyimpangan dari simetri sempurna (Ghasemi & Zahediasl, 2012). Penilaian normalitas sebaiknya mempertimbangkan histogram, Q–Q plot, ukuran sampel, outlier, dan tujuan analisis (Ghasemi & Zahediasl, 2012). Ketergantungan berlebihan pada satu uji normalitas dapat menyesatkan karena uji normalitas dipengaruhi oleh ukuran sampel (Ghasemi & Zahediasl, 2012).

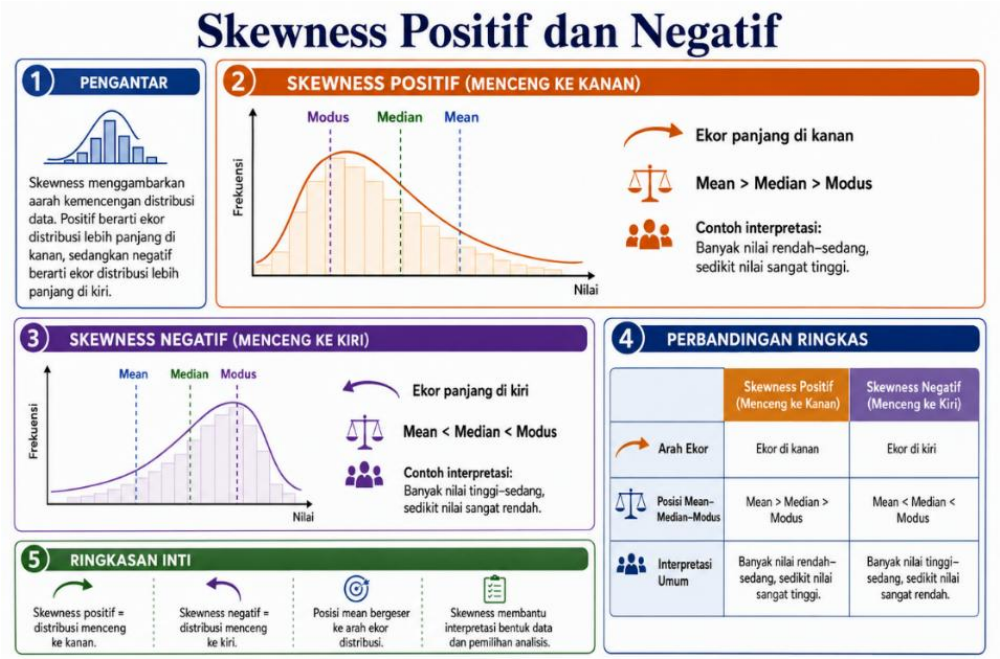


Gambar 4. 3. Distribusi normal dan posisi mean–median–modus
Gambar menunjukkan kurva normal berbentuk lonceng, sebaran simetris, posisi mean, median, dan modus yang berdekatan, serta area sekitar simpangan baku. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012) serta Kamath *et al*. (2025).

4.4 Distribusi Tidak Normal dan Skewness

Distribusi tidak normal adalah distribusi yang menyimpang dari bentuk simetris distribusi normal (Ghasemi & Zahediasl, 2012). Salah satu bentuk distribusi tidak normal yang paling sering ditemukan adalah distribusi menceng atau skewed distribution (Ghasemi & Zahediasl, 2012). Skewness positif atau menceng ke kanan terjadi ketika ekor distribusi memanjang ke arah nilai besar (Ghasemi & Zahediasl, 2012). Skewness negatif atau menceng ke kiri terjadi ketika ekor distribusi memanjang ke arah nilai kecil (Ghasemi & Zahediasl, 2012).

Data klinis seperti lama rawat inap, biaya kesehatan, kadar biomarker tertentu, dan jumlah kunjungan sering dapat menceng ke kanan (Cooksey, 2020). Pada distribusi menceng ke kanan, mean biasanya lebih besar daripada median karena mean tertarik oleh nilai ekstrem di sisi kanan (Cooksey, 2020). Pada distribusi menceng, median dan interquartile range sering lebih representatif daripada mean dan simpangan baku (Cooksey, 2020). Pemilihan uji statistik pada distribusi menceng dapat melibatkan transformasi data, uji nonparametrik, atau model statistik yang sesuai dengan bentuk distribusi (Chicco *et al*., 2025).



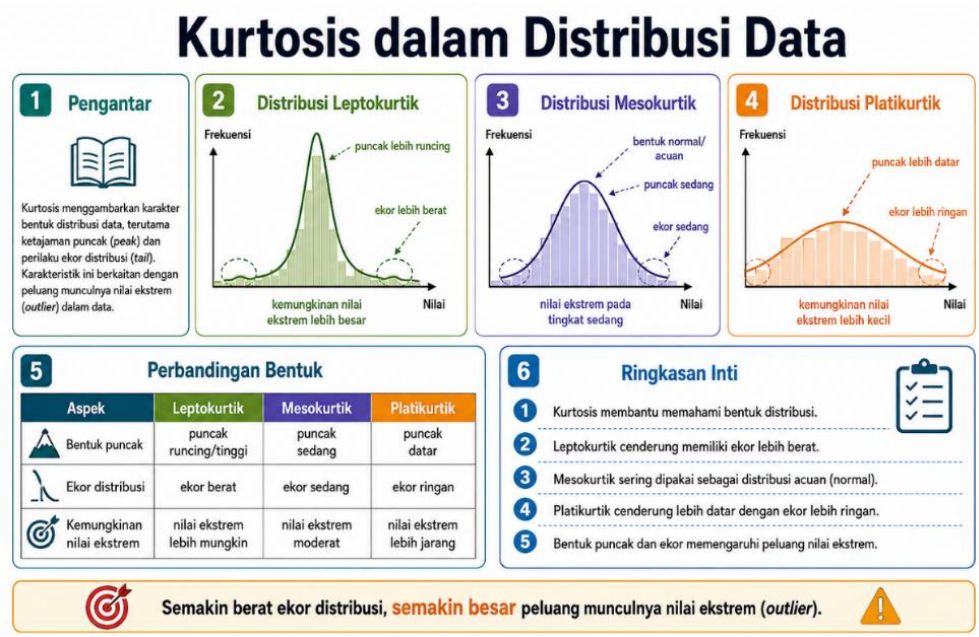
Gambar 4. 4. Skewness positif dan negative
Diagram menunjukkan distribusi menceng ke kanan dengan mean lebih besar dari median dan distribusi menceng ke kiri dengan mean lebih kecil dari median. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012), Cooksey (2020), serta Chicco *et al*. (2025).

4.5 Kurtosis dan Bentuk Puncak Distribusi

Kurtosis menggambarkan bentuk puncak dan ketebalan ekor distribusi data (Ghasemi & Zahediasl, 2012). Distribusi dengan kurtosis tinggi memiliki puncak lebih tajam dan ekor lebih berat dibanding distribusi normal (Ghasemi & Zahediasl, 2012). Distribusi dengan kurtosis rendah memiliki puncak lebih datar dan ekor lebih ringan dibanding distribusi normal (Ghasemi & Zahediasl, 2012). Kurtosis

penting karena ekor distribusi yang berat dapat meningkatkan kemungkinan nilai ekstrem atau outlier (Ghasemi & Zahediasl, 2012).

Penilaian kurtosis sebaiknya tidak dilakukan secara terpisah dari histogram, boxplot, dan pemeriksaan outlier (Ghasemi & Zahediasl, 2012). Data dengan ekor berat dapat memengaruhi simpangan baku dan uji parametrik karena nilai ekstrem memiliki pengaruh besar terhadap estimasi berbasis mean (Ghasemi & Zahediasl, 2012). Jika data menunjukkan kurtosis ekstrem, peneliti dapat mempertimbangkan median, interquartile range, transformasi data, uji nonparametrik, atau statistik robust (Chicco *et al*., 2025). Dengan demikian, kurtosis membantu peneliti memahami apakah distribusi data memiliki konsentrasi ekstrem yang dapat mengganggu analisis (Ghasemi & Zahediasl, 2012).



Gambar 4. 5. Kurtosis dalam distribusi data

Gambar menunjukkan distribusi leptokurtik, mesokurtik, dan platikurtik serta hubungan bentuk puncak dan ekor distribusi dengan kemungkinan nilai ekstrem. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012) serta Chicco *et al*. (2025).

4.6 Outlier dan Pengaruhnya terhadap Analisis

Outlier adalah nilai observasi yang sangat berbeda dari pola umum data (Ghasemi & Zahediasl, 2012). Outlier dapat muncul karena variasi biologis yang benar, kesalahan pengukuran, kesalahan entri data,

kesalahan satuan, atau perbedaan karakteristik subjek (Ghasemi & Zahediasl, 2012). Outlier harus diperiksa sebelum analisis karena dapat menggeser mean, memperbesar simpangan baku, memengaruhi korelasi, mengubah garis regresi, dan memengaruhi hasil uji parametrik (Ghasemi & Zahediasl, 2012). Boxplot dan scatterplot dapat membantu mendeteksi outlier secara visual (Cooksey, 2020).

Outlier tidak boleh langsung dihapus tanpa alasan metodologis yang jelas (Rovetta *et al*., 2025). Jika outlier berasal dari kesalahan entri atau kesalahan satuan, peneliti perlu memperbaiki data berdasarkan sumber asli (Rovetta *et al*., 2025). Jika outlier merupakan nilai biologis yang valid, peneliti perlu mempertimbangkan analisis robust, transformasi data, analisis sensitivitas, atau pelaporan median dan interquartile range (Chicco *et al*., 2025). Pengelolaan outlier harus dilaporkan secara transparan karena keputusan terhadap outlier dapat memengaruhi hasil dan interpretasi penelitian (Rovetta *et al*., 2025).



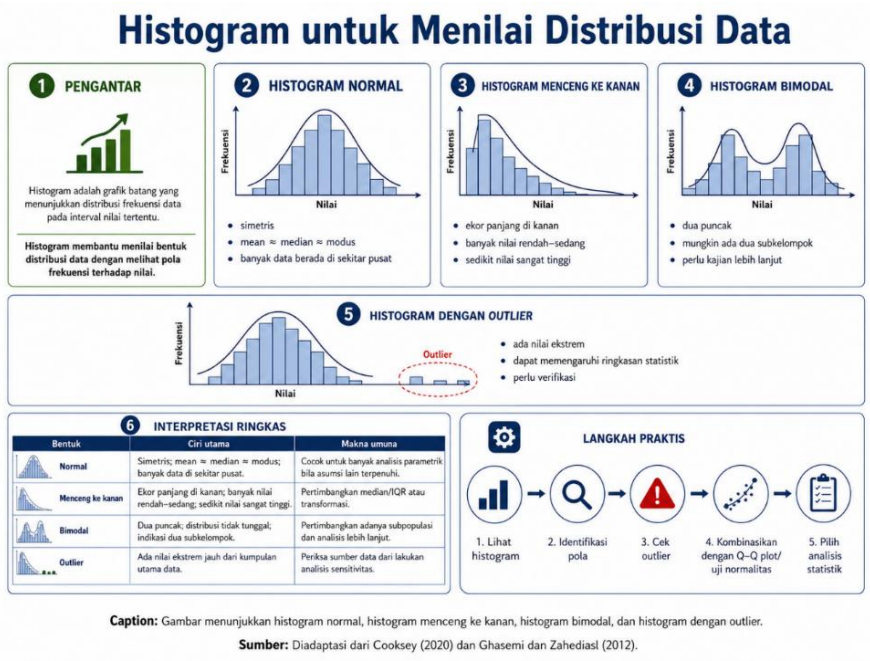
Gambar 4. 6. Outlier dan strategi penanganannya

Diagram menunjukkan outlier akibat kesalahan entri, kesalahan alat ukur, variasi biologis valid, pemeriksaan data sumber, analisis sensitivitas, transformasi, dan statistik robust. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012), Cooksey (2020), Chicco *et al*. (2025), serta Rovetta *et al*. (2025).

4.7 Histogram sebagai Pemeriksaan Distribusi

Histogram adalah grafik yang menunjukkan frekuensi data numerik dalam interval atau bin tertentu (Cooksey, 2020). Histogram membantu peneliti melihat bentuk distribusi, kemencengan, multimodalitas, dan nilai ekstrem (Cooksey, 2020). Histogram juga dapat membantu mengevaluasi apakah data mendekati distribusi normal atau tidak (Ghasemi & Zahediasl, 2012). Interpretasi histogram perlu mempertimbangkan jumlah bin karena terlalu sedikit atau terlalu banyak bin dapat mengubah kesan visual distribusi (Cooksey, 2020).

Histogram sebaiknya digunakan bersama ukuran ringkasan statistik dan grafik lain (Ghasemi & Zahediasl, 2012). Jika histogram menunjukkan distribusi simetris, mean dan simpangan baku dapat menjadi ringkasan yang sesuai (Ghasemi & Zahediasl, 2012). Jika histogram menunjukkan distribusi menceng, median dan interquartile range lebih tepat dipertimbangkan (Cooksey, 2020). Jika histogram menunjukkan beberapa puncak, peneliti perlu memeriksa kemungkinan adanya subkelompok dalam data (Cooksey, 2020).



Caption: Gambar menunjukkan histogram normal, histogram menceng ke kanan, histogram bimodal, dan histogram dengan outlier.
Sumber: Diadaptasi dari Cooksey (2020) dan Ghasemi dan Zahediasl (2012).

Gambar 4. 7. Histogram untuk menilai distribusi data

Gambar menunjukkan histogram normal, histogram menceng ke kanan, histogram bimodal, dan histogram dengan outlier. **Sumber:** Diadaptasi dari Cooksey (2020) dan Ghasemi dan Zahediasl (2012).

4.8 Q–Q Plot dan Pemeriksaan Visual Normalitas

Q–Q plot adalah grafik yang membandingkan kuantil data observasi dengan kuantil distribusi teoretis, biasanya distribusi normal (Ghasemi & Zahediasl, 2012). Jika data mendekati normal, titik-titik pada Q–Q plot cenderung mengikuti garis diagonal (Ghasemi & Zahediasl, 2012). Jika data tidak normal, titik-titik akan menyimpang dari garis diagonal, terutama pada bagian ekor distribusi (Ghasemi & Zahediasl, 2012). Q–Q plot sangat berguna karena dapat memperlihatkan penyimpangan normalitas yang tidak selalu terlihat jelas dari angka ringkasan (Ghasemi & Zahediasl, 2012).

Q–Q plot harus ditafsirkan dengan mempertimbangkan ukuran sampel (Ghasemi & Zahediasl, 2012). Pada sampel kecil, pola titik dapat tampak tidak stabil karena jumlah observasi terbatas (Ghasemi & Zahediasl, 2012). Pada sampel besar, penyimpangan kecil dari garis diagonal dapat terlihat jelas tetapi belum tentu bermakna secara praktis (Ghasemi & Zahediasl, 2012). Oleh karena itu, Q–Q plot sebaiknya digunakan bersama histogram, boxplot, uji normalitas, dan pengetahuan substantif tentang variabel penelitian (Ghasemi & Zahediasl, 2012).

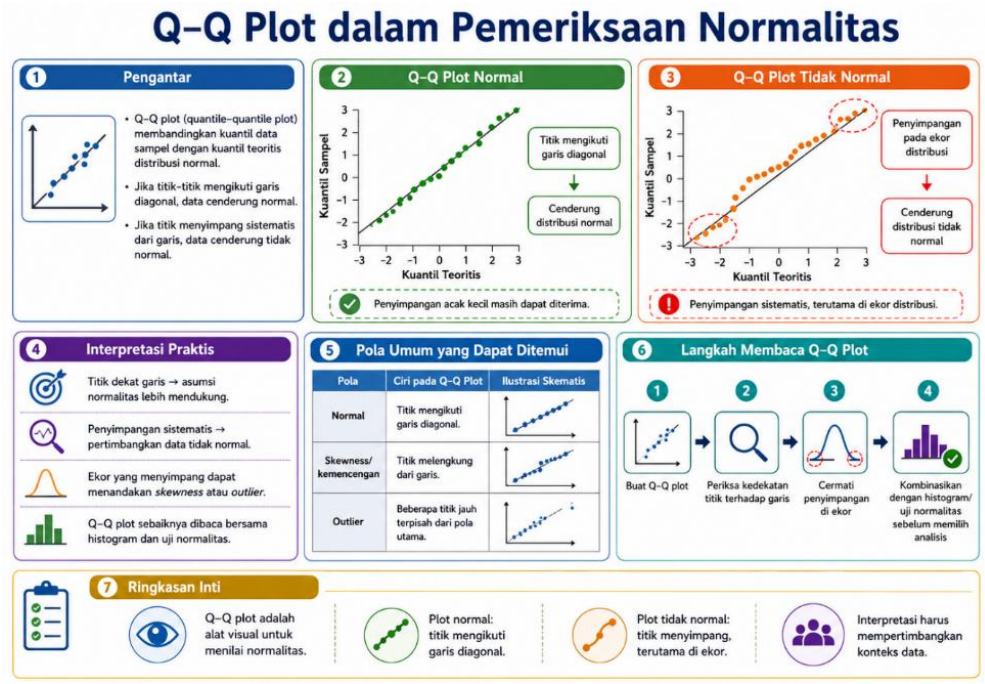
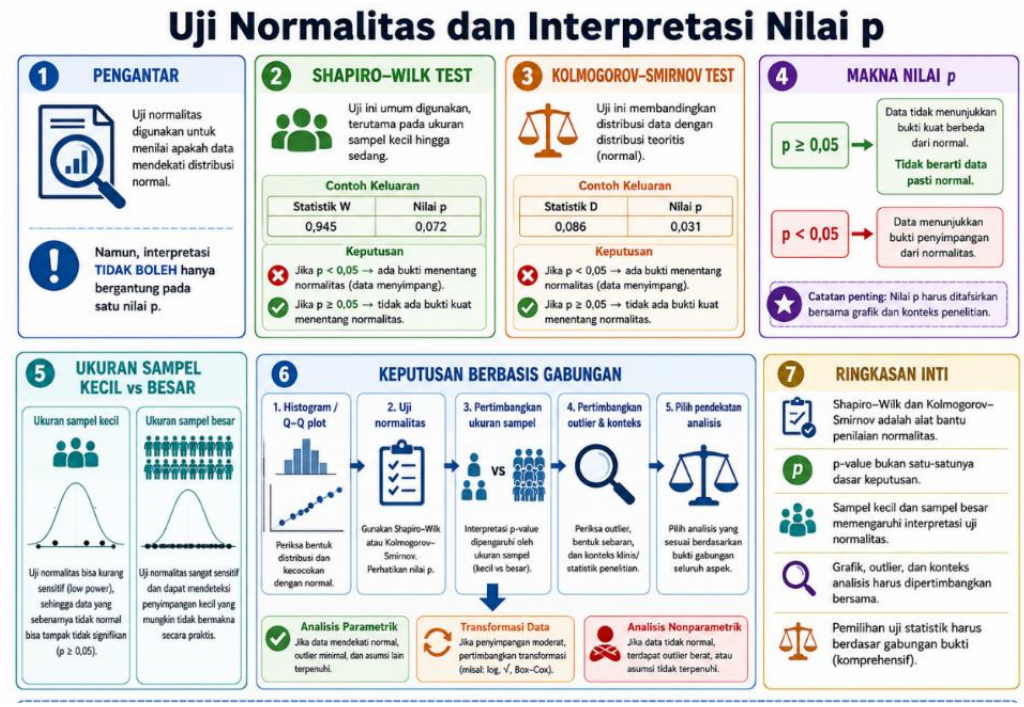


Diagram 4.4. Q–Q plot dalam pemeriksaan normalitas
Caption: Diagram menunjukkan Q–Q plot normal dengan titik mengikuti garis diagonal dan Q–Q plot tidak normal dengan penyimpangan pada ekor distribusi. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012).

4.9 Uji Shapiro–Wilk dan Kolmogorov–Smirnov

Uji Shapiro–Wilk dan Kolmogorov–Smirnov merupakan uji statistik yang sering digunakan untuk menilai normalitas data (Ghasemi & Zahediasl, 2012). Uji Shapiro–Wilk sering direkomendasikan untuk sampel kecil hingga sedang karena memiliki sensitivitas yang baik dalam mendeteksi penyimpangan normalitas (Ghasemi & Zahediasl, 2012). Uji Kolmogorov–Smirnov membandingkan distribusi empiris data dengan distribusi teoretis tertentu (Ghasemi & Zahediasl, 2012). Kedua uji tersebut menghasilkan nilai p yang menunjukkan apakah data berbeda secara statistik dari distribusi normal teoretis (Ghasemi & Zahediasl, 2012).

Uji normalitas memiliki keterbatasan karena sangat dipengaruhi ukuran sampel (Ghasemi & Zahediasl, 2012). Pada sampel kecil, uji normalitas dapat gagal mendeteksi penyimpangan distribusi yang penting (Ghasemi & Zahediasl, 2012). Pada sampel besar, penyimpangan kecil dapat menjadi signifikan secara statistik meskipun pengaruh praktisnya kecil (Ghasemi & Zahediasl, 2012). Oleh karena itu, keputusan memilih uji parametrik atau nonparametrik sebaiknya tidak didasarkan hanya pada satu nilai p dari uji normalitas (Ghasemi & Zahediasl, 2012; Rovetta *et al*., 2025).



Gambar 4.5. Uji normalitas dan interpretasi nilai p
Caption: Gambar menunjukkan Shapiro–Wilk test, Kolmogorov–Smirnov test, nilai p, ukuran sampel kecil, ukuran sampel besar, dan keputusan berbasis gabungan grafik serta konteks. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012) serta Rovetta *et al*. (2025).

4.10 Central Limit Theorem

Central limit theorem menjelaskan bahwa distribusi sampling dari mean cenderung mendekati distribusi normal ketika ukuran sampel cukup besar (Ghasemi & Zahediasl, 2012). Konsep ini menjadi dasar penting dalam banyak prosedur statistik parametrik karena analisis sering berfokus pada distribusi estimasi, bukan hanya distribusi data mentah (Ghasemi & Zahediasl, 2012). Central limit theorem membantu menjelaskan mengapa beberapa uji parametrik tetap cukup robust pada sampel besar meskipun data mentah tidak sepenuhnya normal (Ghasemi & Zahediasl, 2012). Namun, konsep ini tidak berarti bahwa semua data tidak normal dapat langsung dianalisis secara parametrik tanpa pemeriksaan tambahan (Ghasemi & Zahediasl, 2012).

Central limit theorem memiliki batas praktis yang perlu diperhatikan (Ghasemi & Zahediasl, 2012). Jika data sangat menceng, memiliki outlier ekstrem, atau ukuran sampel per kelompok kecil, distribusi sampling dapat tetap menyimpang dari normalitas (Ghasemi &

Zahediasl, 2012). Jika analisis menggunakan model regresi, normalitas residual lebih relevan daripada normalitas variabel mentah pada banyak konteks (Ghasemi & Zahediasl, 2012). Oleh karena itu, central limit theorem harus digunakan sebagai prinsip penalaran statistik, bukan sebagai alasan untuk mengabaikan pemeriksaan distribusi dan asumsi model (Rovetta *et al*., 2025).

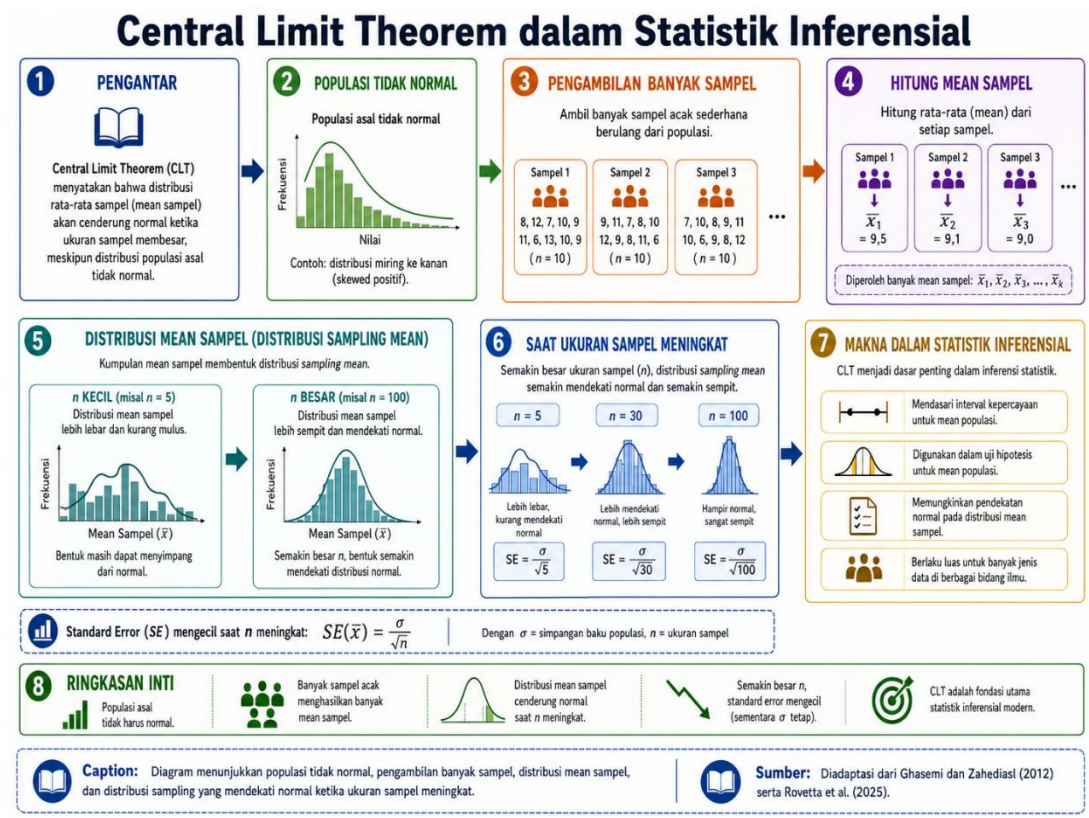


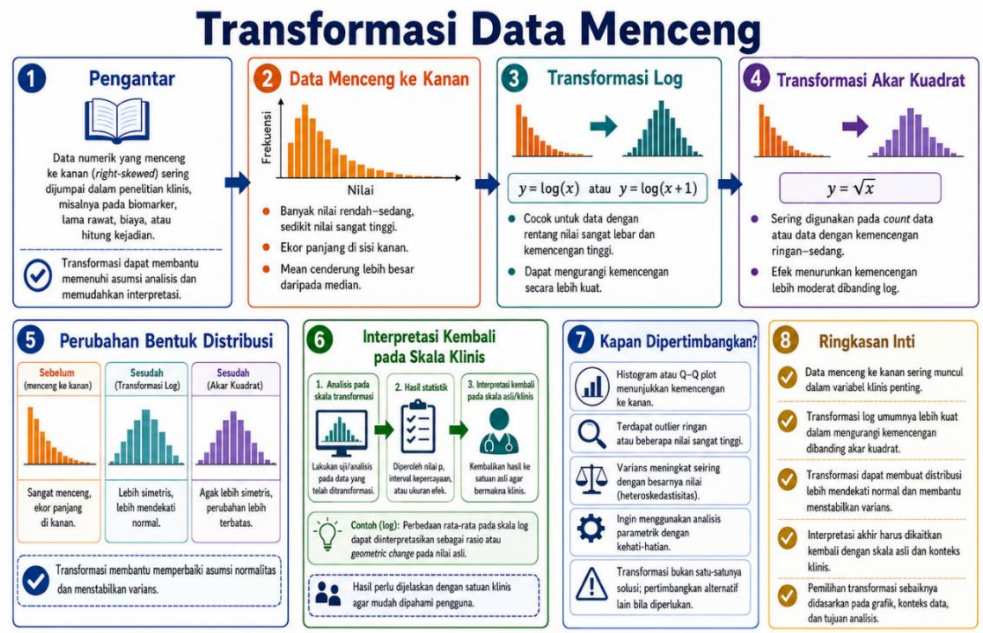
Diagram 4.5. Central limit theorem dalam statistik inferensial
Caption: Diagram menunjukkan populasi tidak normal, pengambilan banyak sampel, distribusi mean sampel, dan distribusi sampling yang mendekati normal ketika ukuran sampel meningkat. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012) serta Rovetta *et al*. (2025).

4.11 Transformasi Data

Transformasi data adalah proses mengubah skala data untuk mengurangi kemencengan, menstabilkan varians, atau membantu memenuhi asumsi model statistik (Ghasemi & Zahediasl, 2012). Transformasi logaritmik sering digunakan pada data yang menceng ke kanan, seperti biaya, konsentrasi biomarker, atau waktu kejadian tertentu

(Ghasemi & Zahediasl, 2012). Transformasi akar kuadrat dapat digunakan pada beberapa jenis data hitungan atau data dengan variasi yang meningkat bersama mean (Ghasemi & Zahediasl, 2012). Transformasi reciprocal kadang digunakan untuk data yang sangat menceng, tetapi interpretasinya dapat menjadi lebih sulit (Ghasemi & Zahediasl, 2012).

Transformasi data harus dilakukan dengan alasan metodologis yang jelas dan sebaiknya direncanakan sebelum analisis utama (Rovetta *et al*., 2025). Setelah data ditransformasi, interpretasi hasil harus memperhatikan skala transformasi yang digunakan (Rovetta *et al*., 2025). Peneliti dapat melaporkan hasil pada skala asli dan skala transformasi jika hal tersebut membantu interpretasi klinis (Rovetta *et al*., 2025). Jika transformasi tidak memperbaiki asumsi atau membuat interpretasi terlalu sulit, uji nonparametrik atau model statistik alternatif dapat dipertimbangkan (Chicco *et al*., 2025).



Gambar 4. 8. Transformasi data menceng

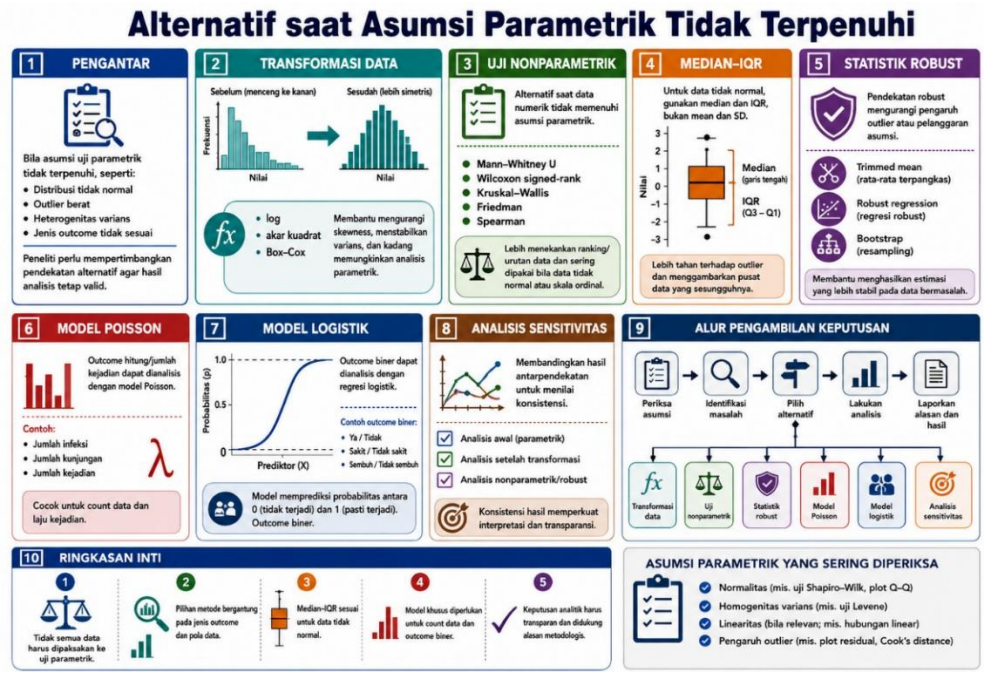
Gambar menunjukkan data menceng ke kanan, transformasi log, transformasi akar kuadrat, perubahan bentuk distribusi, dan interpretasi kembali pada skala klinis. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012), Chicco *et al*. (2025), serta Rovetta *et al*. (2025).

4.12 Statistik Robust dan Alternatif Analisis

Statistik robust adalah pendekatan statistik yang bertujuan mengurangi pengaruh outlier, distribusi tidak normal, atau pelanggaran

asumsi tertentu terhadap hasil analisis (Chicco *et al*., 2025). Median dan interquartile range dapat dianggap lebih robust daripada mean dan simpangan baku pada data yang menceng atau mengandung outlier (Cooksey, 2020). Uji nonparametrik seperti Mann–Whitney U test, Wilcoxon signed-rank test, dan Kruskal–Wallis test dapat menjadi alternatif ketika data tidak memenuhi asumsi parametrik (Chicco *et al*., 2025). Model statistik tertentu juga dapat dipilih sesuai distribusi outcome, seperti model Poisson untuk count data atau model logistik untuk outcome biner (Kosin Medical Journal, 2025).

Pendekatan robust tidak berarti peneliti dapat mengabaikan kualitas data (Rovetta *et al*., 2025). Peneliti tetap harus memeriksa kesalahan entri, missing data, outlier, dan kesesuaian desain analisis (Rovetta *et al*., 2025). Pilihan analisis robust harus didasarkan pada pertanyaan penelitian, jenis outcome, distribusi data, dan tujuan interpretasi (Chicco *et al*., 2025). Dengan demikian, statistik robust merupakan strategi untuk meningkatkan ketahanan analisis, bukan pengganti perencanaan data yang baik (Rovetta *et al*., 2025).



Gambar 4. 9. Alternatif saat asumsi parametrik tidak terpenuhi

Caption: Diagram menunjukkan transformasi data, uji nonparametrik, median-IQR, statistik robust, model Poisson, model logistik, dan analisis sensitivitas. **Sumber:** Diadaptasi dari Chicco *et al*. (2025), Kosin Medical Journal (2025), Cooksey (2020), serta Rovetta *et al*. (2025).

4.13 Pemilihan Uji Parametrik dan Nonparametrik Berdasarkan Distribusi

Pemilihan uji parametrik atau nonparametrik harus mempertimbangkan jenis variabel, distribusi data, ukuran sampel, desain penelitian, dan tujuan analisis (Chicco *et al*., 2025). Independent t-test dapat digunakan untuk membandingkan dua kelompok independen jika outcome numerik memenuhi asumsi parametrik (Chicco *et al*., 2025). Mann–Whitney U test dapat digunakan sebagai alternatif ketika outcome bersifat ordinal atau distribusi numerik tidak sesuai untuk uji parametrik (Chicco *et al*., 2025). Paired t-test digunakan untuk dua pengukuran berpasangan dengan asumsi parametrik yang sesuai, sedangkan Wilcoxon signed-rank test digunakan untuk data berpasangan yang tidak memenuhi asumsi tersebut (Chicco *et al*., 2025).

One-way ANOVA digunakan untuk membandingkan lebih dari dua kelompok independen pada outcome numerik yang memenuhi asumsi parametrik (Chicco *et al*., 2025). Kruskal–Wallis test digunakan untuk membandingkan lebih dari dua kelompok independen ketika outcome ordinal atau tidak memenuhi asumsi parametrik (Chicco *et al*., 2025). Pearson correlation digunakan untuk hubungan linear antara dua variabel numerik yang memenuhi asumsi tertentu, sedangkan Spearman correlation digunakan untuk data ordinal atau hubungan monotonic yang tidak memerlukan normalitas ketat (Chicco *et al*., 2025). Chi-square test dan Fisher's exact test digunakan untuk data kategorik, sehingga tidak diklasifikasikan sebagai uji parametrik atau nonparametrik berdasarkan normalitas data numerik (Kosin Medical Journal, 2025).

Tabel 4. 1. Pemilihan uji berdasarkan distribusi dan desain penelitian

Tujuan Analisis	Kondisi Parametrik	Alternatif Nonparametrik
Dua kelompok independen	Independent t-test	Mann–Whitney U test
Dua pengukuran berpasangan	Paired t-test	Wilcoxon signed-rank test
Lebih dari dua kelompok independen	One-way ANOVA	Kruskal–Wallis test
Hubungan dua variabel numerik	Pearson correlation	Spearman correlation

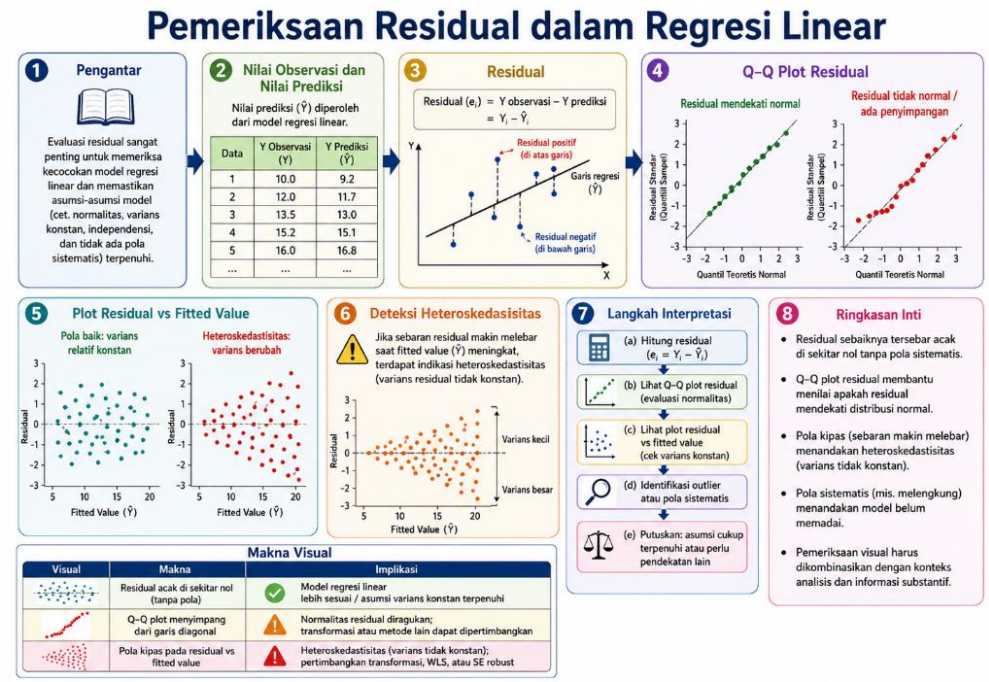
Tujuan Analisis	Kondisi Parametrik	Alternatif Nonparametrik
Data kategorik antar kelompok	Chi-square test	Fisher's exact test bila expected count rendah

Sumber: Diadaptasi dari Chicco *et al*. (2025), Kosin Medical Journal (2025), dan Ghasemi dan Zahediasl (2012).

4.14 Normalitas pada Regresi dan Residual

Dalam analisis regresi linear, asumsi normalitas sering lebih relevan pada residual daripada pada variabel outcome mentah (Ghasemi & Zahediasl, 2012). Residual adalah selisih antara nilai observasi dan nilai prediksi dari model regresi (Ghasemi & Zahediasl, 2012). Pemeriksaan residual membantu peneliti menilai apakah model regresi sesuai dengan pola data (Ghasemi & Zahediasl, 2012). Selain normalitas residual, regresi linear juga memerlukan perhatian terhadap linearitas, homoskedastisitas, independensi observasi, dan outlier berpengaruh (Kosin Medical Journal, 2025).

Jika residual tidak normal atau menunjukkan pola yang sistematis, peneliti perlu mengevaluasi ulang model (Kosin Medical Journal, 2025). Alternatif yang dapat dipertimbangkan adalah transformasi outcome, penambahan bentuk nonlinear, regresi robust, generalized linear model, atau model lain yang sesuai dengan distribusi outcome (Kosin Medical Journal, 2025). Untuk outcome biner, regresi logistik lebih sesuai daripada regresi linear biasa (Kosin Medical Journal, 2025). Untuk data waktu-ke-kejadian, analisis survival seperti Cox regression dapat digunakan sesuai tujuan penelitian (Kosin Medical Journal, 2025).



Gambar 4. 10. Pemeriksaan residual dalam regresi linear

Gambar menunjukkan nilai observasi, nilai prediksi, residual, Q–Q plot residual, plot residual terhadap fitted value, dan deteksi pola heteroskedastisitas. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012) serta Kosin Medical Journal (2025).

4.15 Kesalahan Umum dalam Menilai Normalitas

Kesalahan umum pertama adalah menganggap bahwa semua variabel harus berdistribusi normal sebelum analisis dapat dilakukan (Ghasemi & Zahediasl, 2012). Dalam banyak analisis, yang lebih penting adalah distribusi sampling estimasi atau distribusi residual, bukan selalu distribusi mentah setiap variabel (Ghasemi & Zahediasl, 2012). Kesalahan umum kedua adalah menggunakan uji normalitas sebagai satu-satunya dasar untuk memilih uji statistik (Ghasemi & Zahediasl, 2012). Kesalahan umum ketiga adalah mengabaikan histogram dan Q–Q plot padahal grafik sering memberi informasi penting tentang bentuk distribusi dan outlier (Ghasemi & Zahediasl, 2012).

Kesalahan umum keempat adalah langsung menghapus outlier tanpa memeriksa penyebabnya (Rovetta *et al*., 2025). Kesalahan umum kelima adalah memilih uji nonparametrik untuk semua data yang tidak normal tanpa mempertimbangkan ukuran sampel, effect size, dan tujuan estimasi (Chicco *et al*., 2025). Kesalahan umum keenam adalah menganggap hasil signifikan pada uji normalitas selalu berarti uji

parametrik tidak boleh digunakan (Ghasemi & Zahediasl, 2012). Dengan demikian, penilaian normalitas harus dilakukan secara komprehensif dengan menggabungkan grafik, uji statistik, ukuran sampel, teori statistik, dan konteks penelitian (Ghasemi & Zahediasl, 2012).

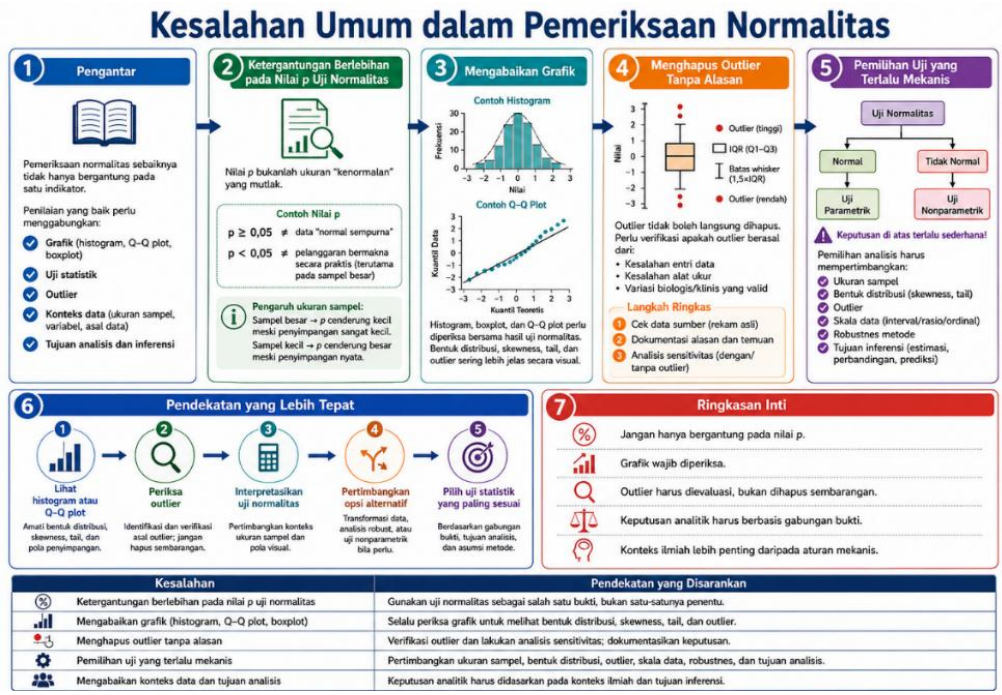


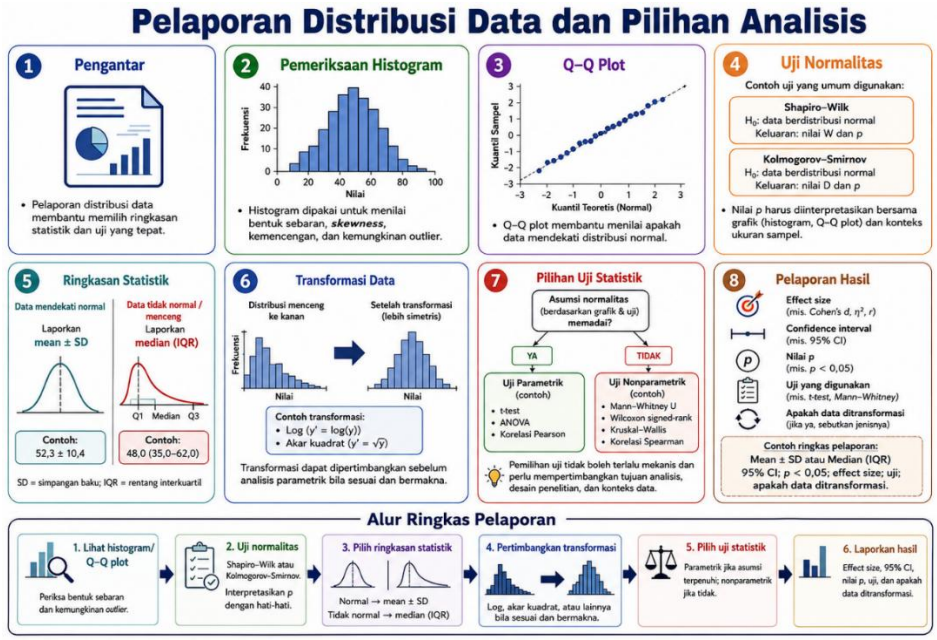
Diagram 4.7. Kesalahan umum dalam pemeriksaan normalitas
Caption: Diagram menunjukkan ketergantungan berlebihan pada nilai p uji normalitas, pengabaian grafik, penghapusan outlier tanpa alasan, dan pemilihan uji yang terlalu mekanis. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012), Chicco *et al*. (2025), serta Rovetta *et al*. (2025).

4.16 Pelaporan Distribusi Data dalam Artikel Ilmiah

Pelaporan distribusi data harus menjelaskan bagaimana data diperiksa dan bagaimana hasil pemeriksaan memengaruhi pilihan analisis statistik (Rovetta *et al*., 2025). Peneliti perlu melaporkan apakah data numerik disajikan sebagai mean-simpangan baku atau median-interquartile range sesuai distribusinya (Cooksey, 2020). Peneliti juga perlu menjelaskan uji normalitas atau grafik yang digunakan bila normalitas menjadi dasar pemilihan analisis (Ghasemi & Zahediasl, 2012). Jika transformasi data dilakukan, jenis transformasi dan alasan transformasi harus dilaporkan (Rovetta *et al*., 2025).

Pelaporan yang baik juga harus menyertakan ukuran efek, confidence interval, dan nilai p bila analisis inferensial dilakukan (Rovetta

et al., 2025). Pelaporan hanya berdasarkan "p < 0,05" tanpa ukuran efek dan interval ketidakpastian dapat menyebabkan interpretasi yang terlalu sederhana (Rovetta *et al*., 2025). Jika uji nonparametrik digunakan, peneliti perlu menyatakan alasan penggunaan uji tersebut, seperti data ordinal, distribusi menceng, atau outlier (Chicco *et al*., 2025). Dengan demikian, pelaporan distribusi data harus membantu pembaca menilai apakah pilihan statistik sesuai dengan karakteristik data (Rovetta *et al*., 2025).



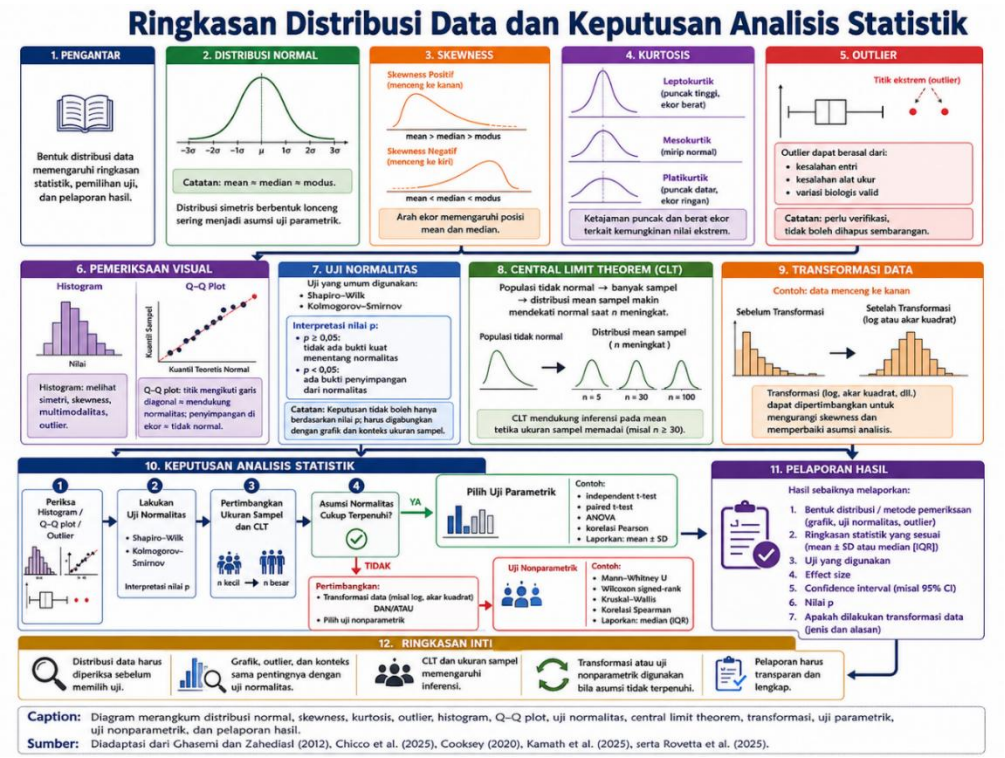
Gambar 4. 11. Pelaporan distribusi data dan pilihan analisis

Gambar menunjukkan pemeriksaan histogram, Q–Q plot, uji normalitas, ringkasan mean-SD, median-IQR, transformasi data, pilihan uji, effect size, confidence interval, dan nilai p. **Sumber:** Diadaptasi dari Rovetta *et al*. (2025), Ghasemi dan Zahediasl (2012), Cooksey (2020), serta Chicco *et al*. (2025).

Bab ini menjelaskan bahwa distribusi data merupakan dasar penting dalam statistik deskriptif dan statistik inferensial (Ghasemi & Zahediasl, 2012). Distribusi normal memiliki bentuk simetris dan menjadi dasar banyak prosedur parametrik, sedangkan distribusi tidak normal dapat menunjukkan skewness, kurtosis ekstrem, multimodalitas, atau outlier (Ghasemi & Zahediasl, 2012). Histogram, boxplot, dan Q–Q plot membantu peneliti memahami bentuk distribusi secara visual (Cooksey, 2020). Uji Shapiro–Wilk dan Kolmogorov–Smirnov dapat digunakan

untuk menguji normalitas, tetapi hasilnya harus ditafsirkan bersama ukuran sampel dan grafik distribusi (Ghasemi & Zahediasl, 2012).

Bab ini juga menegaskan bahwa pemilihan uji parametrik atau nonparametrik tidak boleh dilakukan secara mekanis hanya berdasarkan nilai p uji normalitas (Ghasemi & Zahediasl, 2012). Central limit theorem menjelaskan mengapa beberapa uji parametrik dapat robust pada sampel besar, tetapi outlier ekstrem, distribusi sangat menceng, dan sampel kecil tetap memerlukan perhatian khusus (Ghasemi & Zahediasl, 2012). Transformasi data, statistik robust, uji nonparametrik, dan model alternatif dapat digunakan bila asumsi parametrik tidak sesuai (Chicco *et al*., 2025). Pelaporan hasil analisis harus menyertakan alasan pemilihan uji, ringkasan data yang sesuai, ukuran efek, confidence interval, dan nilai p agar interpretasi ilmiah lebih lengkap (Rovetta *et al*., 2025).



Gambar 4. 12. Ringkasan distribusi data dan keputusan analisis statistic Diagram merangkum distribusi normal, skewness, kurtosis, outlier, histogram, Q–Q plot, uji normalitas, central limit theorem, transformasi, uji parametrik, uji nonparametrik, dan pelaporan hasil. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012), Chicco *et al*. (2025), Cooksey (2020), Kamath *et al*. (2025), serta Rovetta *et al*. (2025).

Review Questions

1. **Apa yang dimaksud distribusi data?**
   Jawaban: Distribusi data adalah pola penyebaran nilai observasi dalam suatu variabel penelitian.
2. **Apa ciri utama distribusi normal?**
   Jawaban: Distribusi normal berbentuk lonceng, simetris, dan memiliki mean, median, serta modus yang berada pada posisi berdekatan.
3. **Apa yang dimaksud skewness positif?**
   Jawaban: Skewness positif adalah distribusi menceng ke kanan, yaitu ekor distribusi memanjang ke arah nilai yang lebih besar.
4. **Apa fungsi histogram dalam analisis statistik?**
   Jawaban: Histogram digunakan untuk melihat bentuk distribusi, kemencengan, multimodalitas, dan kemungkinan outlier pada data numerik.
5. **Apa fungsi Q–Q plot?**
   Jawaban: Q–Q plot digunakan untuk membandingkan kuantil data observasi dengan kuantil distribusi teoretis, terutama distribusi normal.
6. **Mengapa uji normalitas tidak boleh menjadi satu-satunya dasar pemilihan uji statistik?**
   Jawaban: Uji normalitas dipengaruhi oleh ukuran sampel sehingga perlu ditafsirkan bersama histogram, Q–Q plot, outlier, dan konteks penelitian.
7. **Apa yang dimaksud central limit theorem?**
   Jawaban: Central limit theorem menjelaskan bahwa distribusi sampling dari mean cenderung mendekati normal ketika ukuran sampel cukup besar.
8. **Kapan transformasi data dapat dipertimbangkan?**
   Jawaban: Transformasi data dapat dipertimbangkan ketika data sangat menceng, varians tidak stabil, atau asumsi model statistik tidak terpenuhi.
9. **Apa alternatif jika asumsi parametrik tidak terpenuhi?**
   Jawaban: Alternatifnya adalah menggunakan transformasi data, uji nonparametrik, statistik robust, analisis sensitivitas, atau model statistik lain yang sesuai.
10. **Apa yang sebaiknya dilaporkan terkait distribusi data?**
    Jawaban: Peneliti sebaiknya melaporkan cara pemeriksaan distribusi, ringkasan data yang sesuai, alasan pemilihan uji, ukuran efek, confidence interval, dan nilai p.

DAFTAR PUSTAKA

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*, Article 40. https://doi.org/10.1186/s13040-025-00465-6

Cooksey, R. W. (2020). Descriptive statistics for summarising data. In *Illustrating statistical procedures: Finding meaning in quantitative data* (pp. 61–139). Springer.

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

Kamath, A., *et al*. (2025). Assessing the robustness of normality tests under varying sample sizes and distributions in public health research. *PMC-indexed article*.

Kaliyadan, F., & Kulkarni, V. (2019). Types of variables, descriptive statistics, and sample size. *Indian Dermatology Online Journal, 10*(1), 82–86. https://doi.org/10.4103/idoj.IDOJ_468_18

Kosin Medical Journal. (2025). Common statistical methods used in medical research. *Kosin Medical Journal*.

Rovetta, A., Piretta, L., & Mansournia, M. A. (2025). p-Values and confidence intervals as compatibility measures: Guidelines for interpreting statistical studies in clinical research. *The Lancet Regional Health – Southeast Asia*.

BAB 5

UJI PERBEDAAN KELOMPOK: UJI T, MANN–WHITNEY, WILCOXON, ANOVA, DAN KRUSKAL–WALLIS

5.1 Pendahuluan

Uji perbedaan kelompok digunakan ketika peneliti ingin mengetahui apakah nilai suatu outcome berbeda antara dua kelompok atau lebih (Lee, 2022). Pemilihan uji perbedaan kelompok harus didasarkan pada jenis variabel outcome, jumlah kelompok, hubungan antarobservasi, distribusi data, dan asumsi statistik yang digunakan (Chicco *et al*., 2025). Dalam penelitian biomedis dan klinis, uji yang sering digunakan untuk membandingkan kelompok meliputi independent t-test, paired t-test, Mann–Whitney U test, Wilcoxon signed-rank test, one-way ANOVA, dan Kruskal–Wallis test (Chicco *et al*., 2025; Lee, 2022). Uji parametrik seperti t-test dan ANOVA umumnya digunakan untuk outcome numerik yang memenuhi asumsi parametrik, sedangkan uji nonparametrik seperti Mann–Whitney, Wilcoxon, dan Kruskal–Wallis digunakan ketika data ordinal atau asumsi parametrik tidak terpenuhi (Chicco *et al*., 2025).

Bab ini membahas prinsip uji perbedaan dua kelompok independen, dua kelompok berpasangan, lebih dari dua kelompok independen, dan lebih dari dua pengukuran berulang. Bab ini juga menjelaskan peran hipotesis nol, nilai p, confidence interval, effect size, asumsi normalitas, homogenitas varians, dan interpretasi hasil secara ilmiah. Interpretasi hasil uji statistik tidak boleh hanya bergantung pada nilai p, karena nilai p tidak menunjukkan besar efek atau makna klinis secara langsung (Rovetta *et al*., 2025). Oleh karena itu, setiap uji perbedaan kelompok sebaiknya dilaporkan bersama estimasi efek, confidence interval, ukuran efek, dan alasan pemilihan uji (Rovetta *et al*., 2025).



Gambar 5. 1. Peta konsep uji perbedaan kelompok

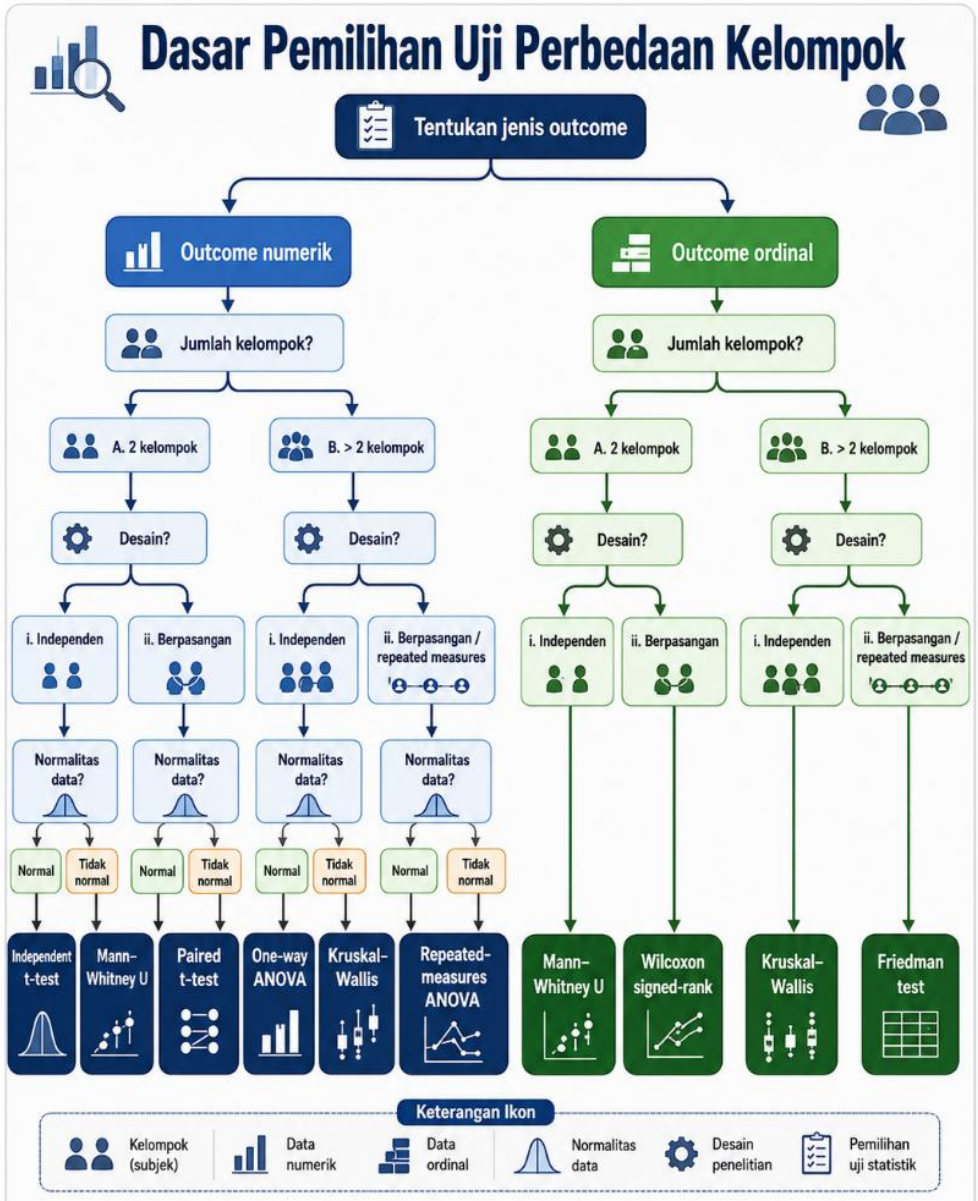
Ilustrasi menunjukkan pemilihan uji berdasarkan jumlah kelompok, desain independen atau berpasangan, distribusi data, dan pilihan uji parametrik atau nonparametrik. **Sumber:** Diadaptasi dari Chicco *et al*. (2025), Lee (2022), Ghasemi dan Zahediasl (2012), serta Rovetta *et al*. (2025).

5.2 Konsep Dasar Uji Perbedaan Kelompok

Uji perbedaan kelompok bertujuan mengevaluasi apakah perbedaan nilai outcome antar kelompok cukup besar dibandingkan variasi data dalam kelompok (Lee, 2022). Pada uji parametrik, perbandingan sering didasarkan pada mean dan variabilitas data, sedangkan pada uji nonparametrik perbandingan sering didasarkan pada peringkat atau distribusi nilai (Chicco *et al*., 2025). Hipotesis nol biasanya menyatakan bahwa tidak ada perbedaan antara kelompok, sedangkan hipotesis alternatif menyatakan bahwa terdapat perbedaan antara kelompok (Rovetta *et al*., 2025). Hasil uji statistik kemudian digunakan untuk menilai kompatibilitas data dengan hipotesis nol berdasarkan asumsi model yang digunakan (Rovetta *et al*., 2025).

Pemilihan uji harus dimulai dari desain penelitian, bukan dari menu perangkat lunak statistik. Jika dua kelompok saling independen, analisis yang digunakan berbeda dari analisis dua pengukuran pada subjek yang sama (Chicco *et al*., 2025). Jika outcome numerik berdistribusi normal dan varians relatif homogen, uji parametrik dapat

digunakan (Ghasemi & Zahediasl, 2012). Jika outcome bersifat ordinal, menceng kuat, atau memiliki outlier ekstrem, uji nonparametrik dapat lebih sesuai (Chicco *et al*., 2025).



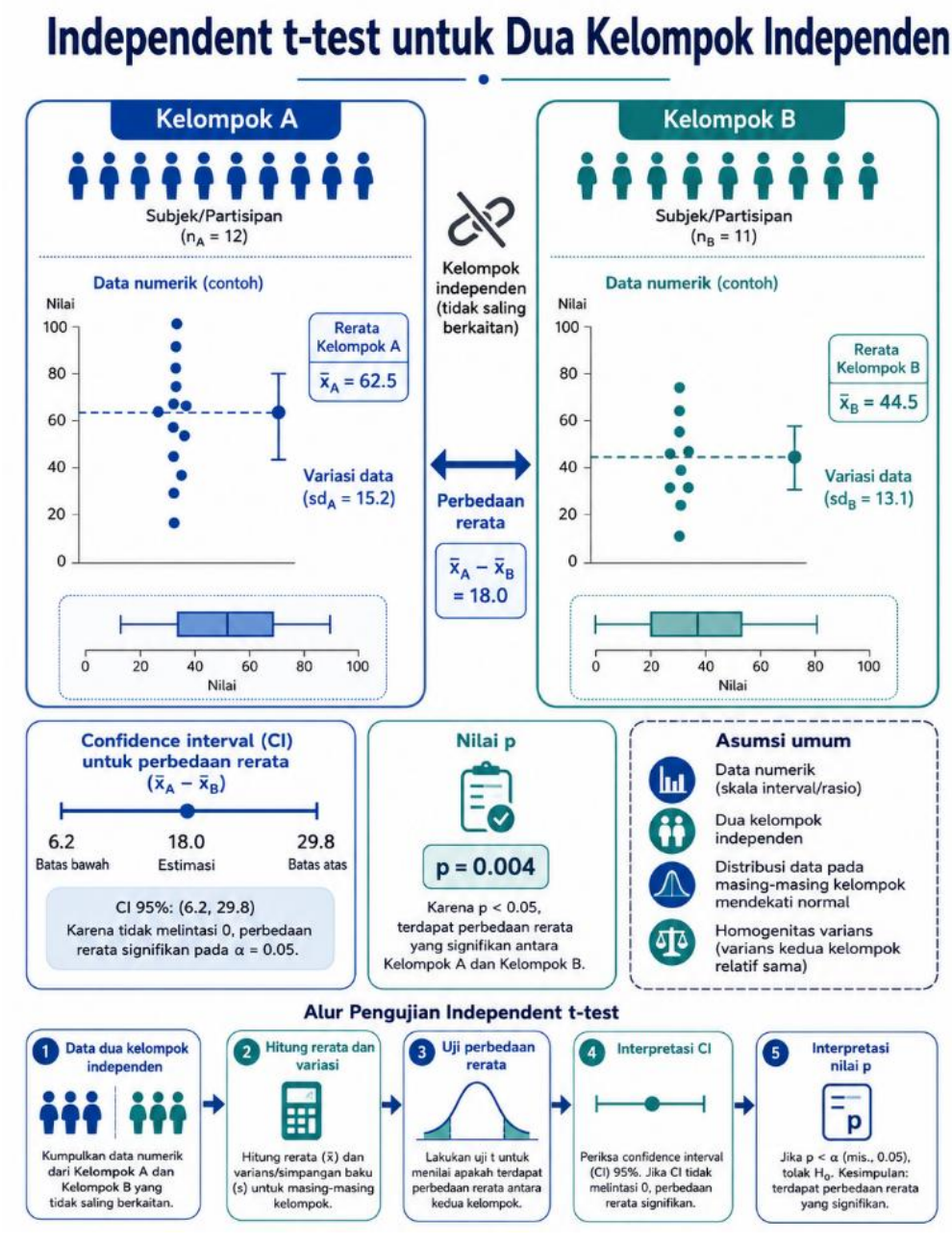
Gambar 5. 2. Dasar pemilihan uji perbedaan kelompok

Diagram menunjukkan hubungan antara outcome numerik atau ordinal, jumlah kelompok, desain independen atau berpasangan, normalitas data, dan pilihan uji statistik. **Sumber:** Diadaptasi dari Chicco *et al*. (2025), Lee (2022), dan Ghasemi dan Zahediasl (2012).

5.3 Independent t-test

Independent t-test digunakan untuk membandingkan rerata outcome numerik antara dua kelompok independen (Chicco *et al*., 2025). Contoh penggunaannya adalah membandingkan rerata tekanan darah antara kelompok terapi dan kontrol atau membandingkan rerata kadar glukosa antara dua kelompok pasien yang berbeda (Lee, 2022). Uji ini termasuk uji parametrik karena analisisnya menggunakan mean, varians, dan asumsi distribusi tertentu (Ghasemi & Zahediasl, 2012). Asumsi penting independent t-test meliputi outcome numerik, observasi independen, distribusi yang mendekati normal dalam tiap kelompok, dan varians yang relatif homogen (Ghasemi & Zahediasl, 2012; Chicco *et al*., 2025).

Jika varians kedua kelompok tidak homogen, Welch's t-test dapat dipertimbangkan sebagai alternatif karena lebih toleran terhadap perbedaan varians antar kelompok (Lee, 2022). Sebelum menggunakan independent t-test, peneliti perlu memeriksa histogram, Q–Q plot, outlier, dan ukuran sampel tiap kelompok (Ghasemi & Zahediasl, 2012). Jika data sangat menceng atau terdapat outlier ekstrem, median dan interquartile range dapat dilaporkan, dan uji nonparametrik seperti Mann–Whitney U test dapat dipertimbangkan (Chicco *et al*., 2025). Hasil independent t-test sebaiknya dilaporkan bersama perbedaan rerata, confidence interval, nilai p, dan ukuran efek seperti Cohen's d (Rovetta *et al*., 2025).

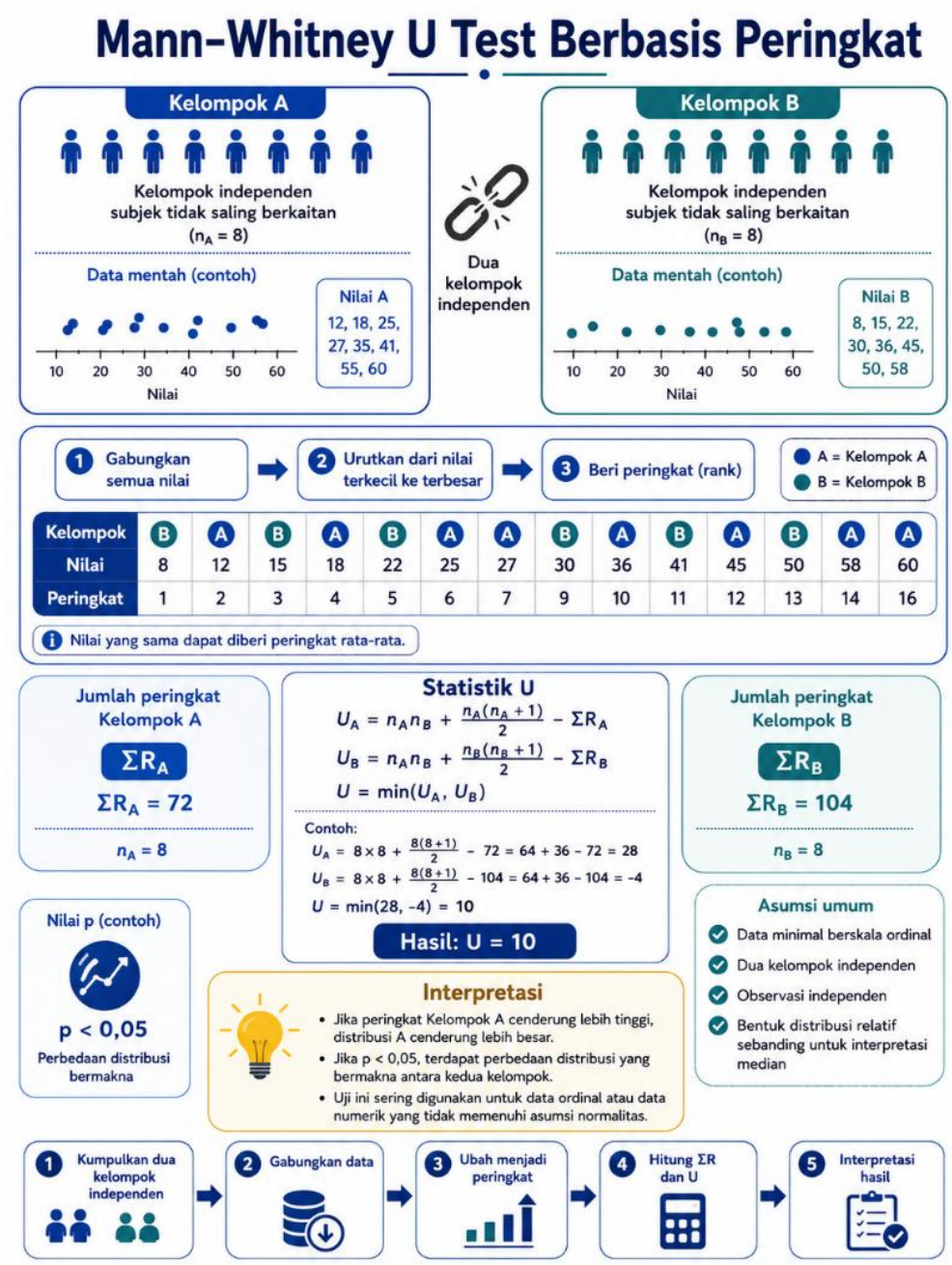


Gambar 5. 3. Independent t-test untuk dua kelompok independent Gambar menunjukkan dua kelompok independen, rerata masing-masing kelompok, variasi data, perbedaan rerata, confidence interval, dan nilai p. **Sumber:** Diadaptasi dari Chicco *et al*. (2025), Lee (2022), Ghasemi dan Zahediasl (2012), serta Rovetta *et al*. (2025).

5.4 Mann–Whitney U Test

Mann–Whitney U test adalah uji nonparametrik untuk membandingkan dua kelompok independen ketika outcome bersifat ordinal atau outcome numerik tidak memenuhi asumsi parametrik (Chicco *et al*., 2025). Uji ini bekerja dengan mengubah nilai observasi menjadi peringkat, kemudian membandingkan distribusi peringkat antar kelompok (Chicco *et al*., 2025). Mann–Whitney U test sering digunakan sebagai alternatif independent t-test ketika data menceng, ukuran sampel kecil, atau terdapat outlier ekstrem (Chicco *et al*., 2025). Namun, interpretasi Mann–Whitney tidak selalu identik dengan "perbedaan median" kecuali bentuk distribusi kedua kelompok serupa (Chicco *et al*., 2025).

Mann–Whitney U test cocok untuk membandingkan skor nyeri antara dua kelompok independen, tingkat kepuasan pasien, atau data numerik yang tidak normal (Chicco *et al*., 2025). Peneliti tetap perlu melaporkan ringkasan data yang sesuai, seperti median dan interquartile range, terutama jika distribusi data tidak normal (Cooksey, 2020). Hasil uji sebaiknya disertai ukuran efek nonparametrik, seperti rank-biserial correlation atau r, bila tersedia (Rovetta *et al*., 2025). Penggunaan Mann–Whitney sebaiknya dijelaskan berdasarkan sifat data dan asumsi analisis, bukan hanya karena hasil uji normalitas signifikan (Ghasemi & Zahediasl, 2012).



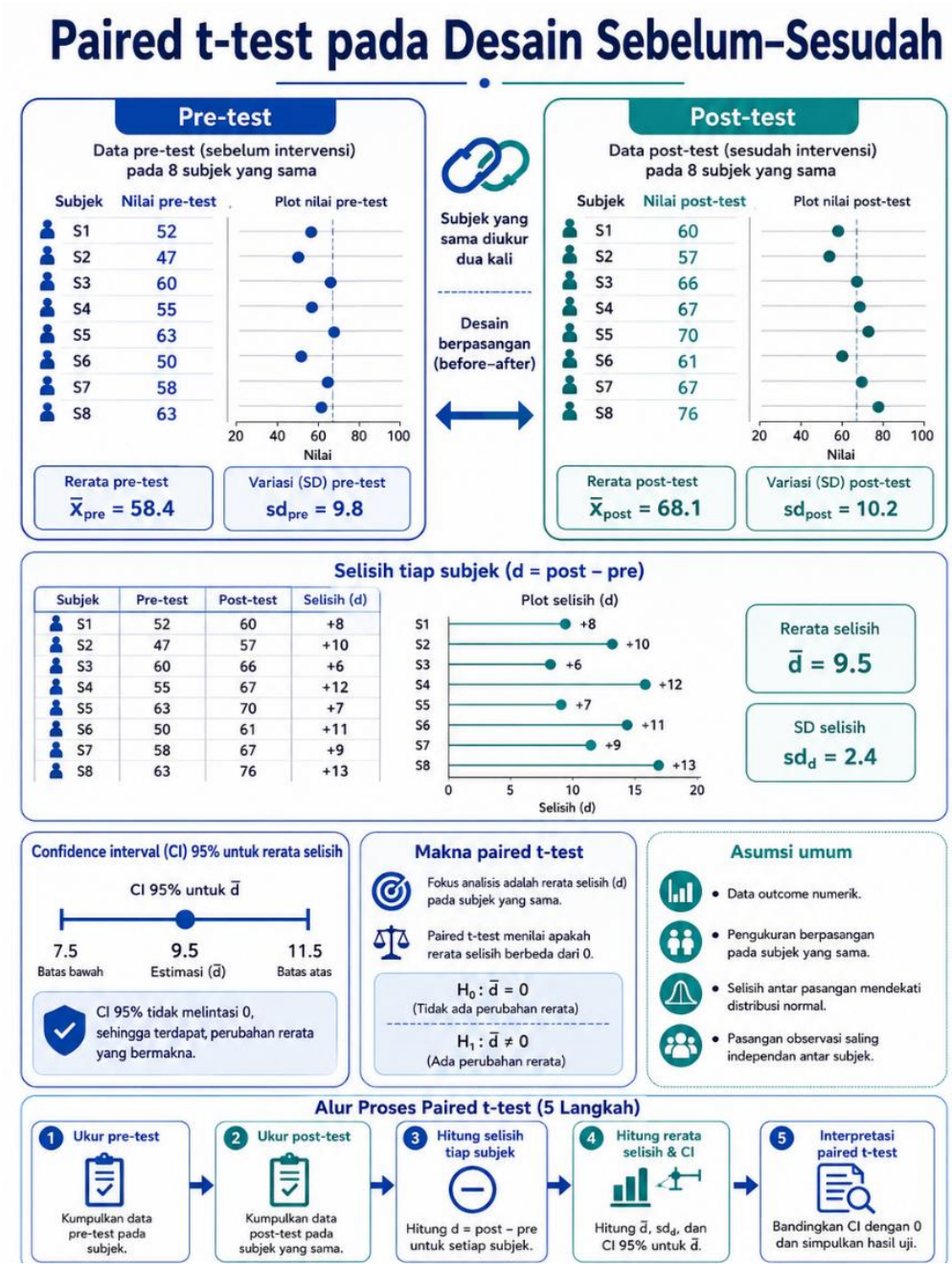
Gambar 5. 4. Mann–Whitney U test berbasis peringkat

Diagram menunjukkan dua kelompok independen, pengurutan nilai menjadi peringkat, jumlah peringkat tiap kelompok, statistik U, dan interpretasi perbedaan distribusi. **Sumber:** Diadaptasi dari Chicco *et al*. (2025), Cooksey (2020), dan Ghasemi dan Zahediasl (2012).

5.5 Paired t-test

Paired t-test digunakan untuk membandingkan rerata dua pengukuran berpasangan pada subjek yang sama atau unit yang dipasangkan (Chicco *et al*., 2025). Contoh penggunaannya adalah membandingkan tekanan darah sebelum dan sesudah terapi pada pasien yang sama atau membandingkan skor nyeri sebelum dan sesudah intervensi (Lee, 2022). Berbeda dari independent t-test, paired t-test menganalisis selisih nilai antar dua pengukuran pada setiap subjek (Chicco *et al*., 2025). Asumsi utama paired t-test adalah selisih antar pasangan berdistribusi mendekati normal, bukan harus setiap variabel mentah berdistribusi normal secara sempurna (Ghasemi & Zahediasl, 2012).

Paired t-test lebih efisien daripada independent t-test untuk desain sebelum–sesudah karena variasi antarindividu dapat dikendalikan melalui pasangan data (Lee, 2022). Peneliti harus memastikan bahwa pasangan data benar-benar berasal dari subjek atau unit yang sama atau dipasangkan secara valid (Chicco *et al*., 2025). Jika selisih antar pasangan sangat menceng atau data bersifat ordinal, Wilcoxon signed-rank test dapat digunakan sebagai alternatif nonparametrik (Chicco *et al*., 2025). Hasil paired t-test sebaiknya dilaporkan sebagai rerata selisih, confidence interval selisih, nilai p, dan ukuran efek (Rovetta *et al*., 2025).



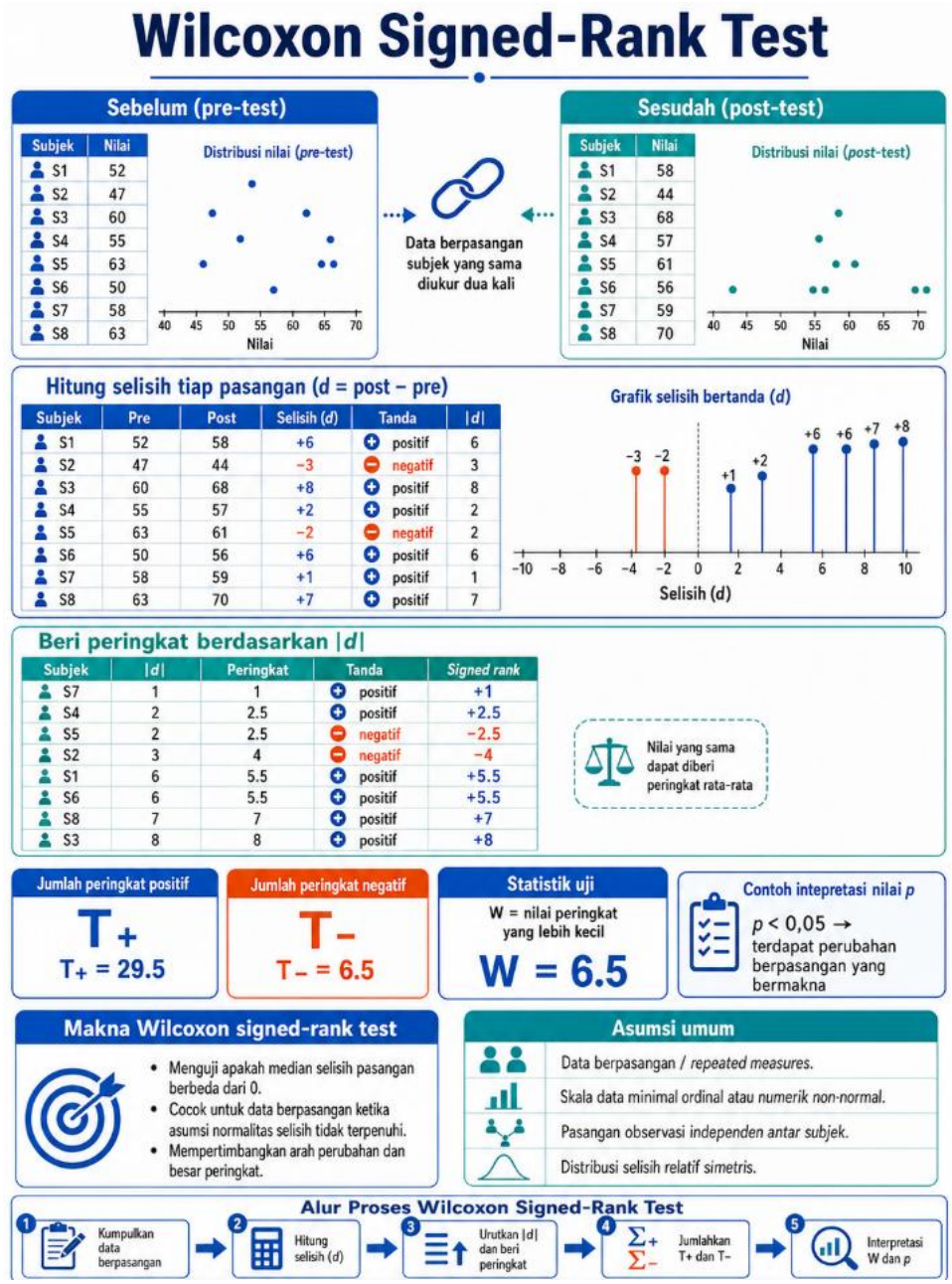
Gambar 5. 5. Paired t-test pada desain sebelum–sesudah

Gambar menunjukkan pengukuran pre-test dan post-test pada subjek yang sama, selisih tiap subjek, rerata selisih, dan confidence interval. **Sumber:** Diadaptasi dari Chicco *et al*. (2025), Lee (2022), Ghasemi dan Zahediasl (2012), serta Rovetta *et al*. (2025).

5.6 Wilcoxon Signed-Rank Test

Wilcoxon signed-rank test adalah uji nonparametrik untuk membandingkan dua pengukuran berpasangan ketika data tidak memenuhi asumsi paired t-test atau outcome bersifat ordinal (Chicco *et al*., 2025). Uji ini menggunakan tanda dan peringkat dari selisih absolut antar pasangan data (Chicco *et al*., 2025). Wilcoxon signed-rank test cocok untuk data sebelum–sesudah yang tidak normal, skor ordinal, atau pengukuran berpasangan dengan outlier yang kuat (Chicco *et al*., 2025). Uji ini berbeda dari sign test karena Wilcoxon mempertimbangkan besar peringkat selisih, bukan hanya arah perubahan (Chicco *et al*., 2025).

Dalam pelaporan Wilcoxon signed-rank test, peneliti sebaiknya menyajikan median dan interquartile range untuk masing-masing waktu pengukuran atau median perubahan jika sesuai (Cooksey, 2020). Peneliti juga perlu menjelaskan bahwa data berasal dari pasangan observasi yang sama, karena uji ini tidak tepat untuk kelompok independen (Chicco *et al*., 2025). Jika banyak selisih bernilai nol, interpretasi perlu lebih hati-hati karena hal tersebut dapat memengaruhi perhitungan peringkat (Chicco *et al*., 2025). Hasil uji sebaiknya dilengkapi dengan ukuran efek dan interpretasi klinis agar tidak bergantung hanya pada nilai p (Rovetta *et al*., 2025).



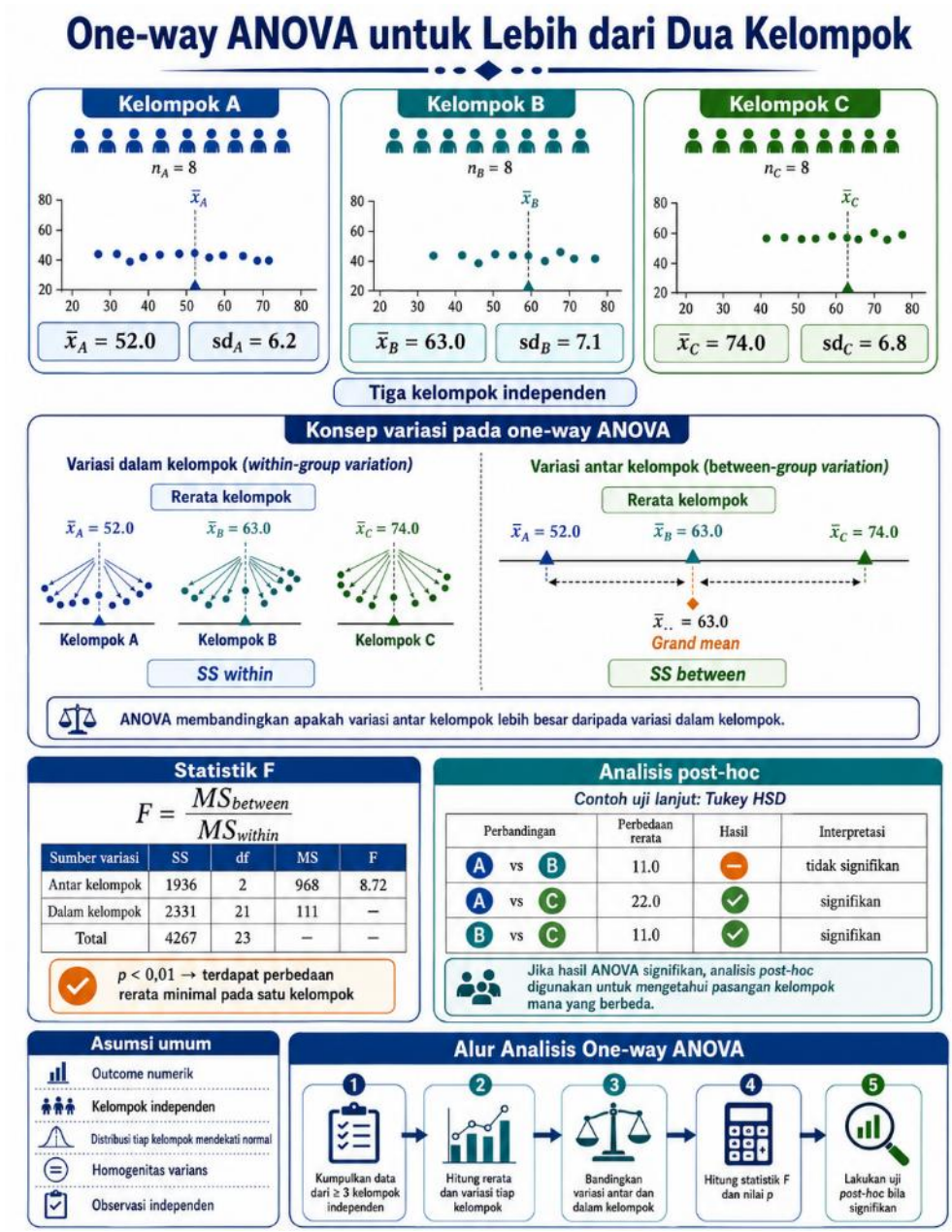
Gambar 5. 6. Wilcoxon signed-rank test

Diagram menunjukkan data berpasangan, perhitungan selisih, pemberian tanda positif-negatif, pengurutan peringkat, dan interpretasi perubahan berpasangan. **Sumber:** Diadaptasi dari Chicco *et al*. (2025), Cooksey (2020), dan Rovetta *et al*. (2025).

5.7 One-Way ANOVA

One-way ANOVA digunakan untuk membandingkan rerata outcome numerik pada lebih dari dua kelompok independen (Lee, 2022). Contoh penggunaannya adalah membandingkan rerata kadar kolesterol pada tiga kelompok diet atau membandingkan rerata skor kualitas hidup pada tiga kelompok terapi (Lee, 2022). ANOVA mengevaluasi apakah variasi antar kelompok lebih besar dibanding variasi dalam kelompok (Lee, 2022). Asumsi one-way ANOVA meliputi outcome numerik, observasi independen, distribusi residual yang mendekati normal, dan homogenitas varians antar kelompok (Ghasemi & Zahediasl, 2012; Lee, 2022).

Jika ANOVA menunjukkan hasil signifikan, peneliti perlu melakukan analisis post-hoc untuk mengetahui kelompok mana yang berbeda (Lee, 2022). Analisis post-hoc harus memperhitungkan risiko inflasi kesalahan tipe I karena banyaknya perbandingan antar kelompok (Lee, 2022). Jika asumsi homogenitas varians tidak terpenuhi, alternatif seperti Welch ANOVA dapat dipertimbangkan sesuai konteks analisis (Lee, 2022). Hasil ANOVA sebaiknya dilaporkan bersama rerata tiap kelompok, simpangan baku, nilai F, derajat kebebasan, nilai p, ukuran efek seperti eta squared, dan hasil post-hoc bila dilakukan (Rovetta *et al*., 2025).



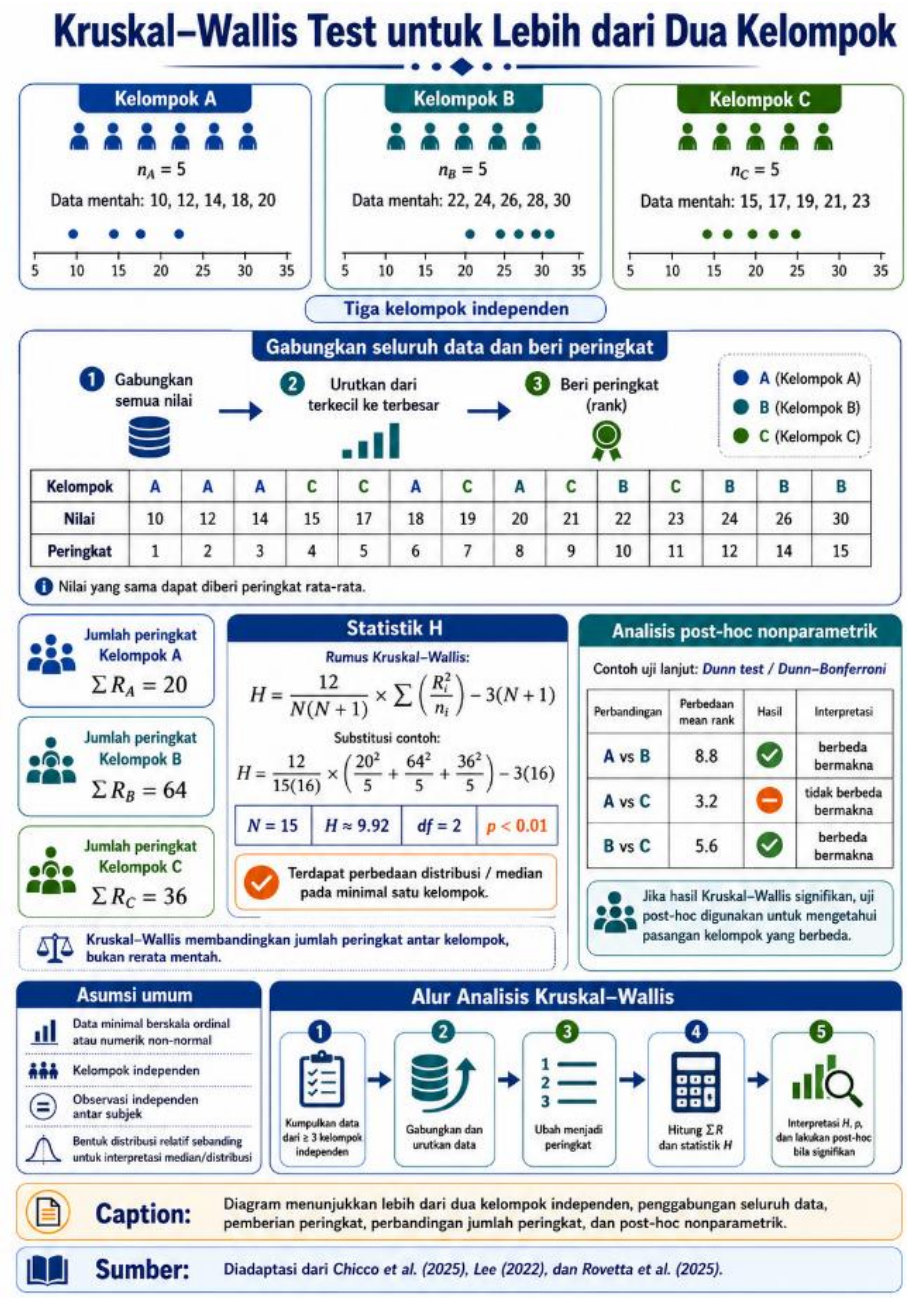
Gambar 5. 7. One-way ANOVA untuk lebih dari dua kelompok

Gambar menunjukkan tiga kelompok independen, rerata tiap kelompok, variasi dalam kelompok, variasi antar kelompok, statistik F, dan analisis post-hoc. **Sumber:** Diadaptasi dari Lee (2022), Ghasemi dan Zahediasl (2012), serta Rovetta *et al*. (2025).

5.8 Kruskal–Wallis Test

Kruskal–Wallis test adalah uji nonparametrik untuk membandingkan lebih dari dua kelompok independen ketika outcome bersifat ordinal atau outcome numerik tidak memenuhi asumsi ANOVA (Chicco *et al*., 2025). Uji ini merupakan perluasan dari Mann–Whitney U test untuk lebih dari dua kelompok independen (Chicco *et al*., 2025). Kruskal–Wallis menggunakan peringkat seluruh observasi, kemudian membandingkan distribusi peringkat antar kelompok (Chicco *et al*., 2025). Uji ini sering digunakan ketika data menceng, memiliki outlier, atau ukuran sampel kelompok relatif kecil (Chicco *et al*., 2025).

Jika Kruskal–Wallis test signifikan, peneliti perlu melakukan uji post-hoc nonparametrik untuk mengetahui pasangan kelompok yang berbeda (Chicco *et al*., 2025). Post-hoc nonparametrik perlu memperhatikan koreksi multiple comparisons untuk mengurangi risiko kesalahan tipe I (Lee, 2022). Pelaporan Kruskal–Wallis sebaiknya menyertakan median, interquartile range, nilai statistik uji, derajat kebebasan, nilai p, dan ukuran efek bila tersedia (Rovetta *et al*., 2025). Interpretasi hasil Kruskal–Wallis harus memperhatikan apakah distribusi antar kelompok memiliki bentuk yang serupa, karena perbedaan peringkat tidak selalu berarti perbedaan median sederhana (Chicco *et al*., 2025).



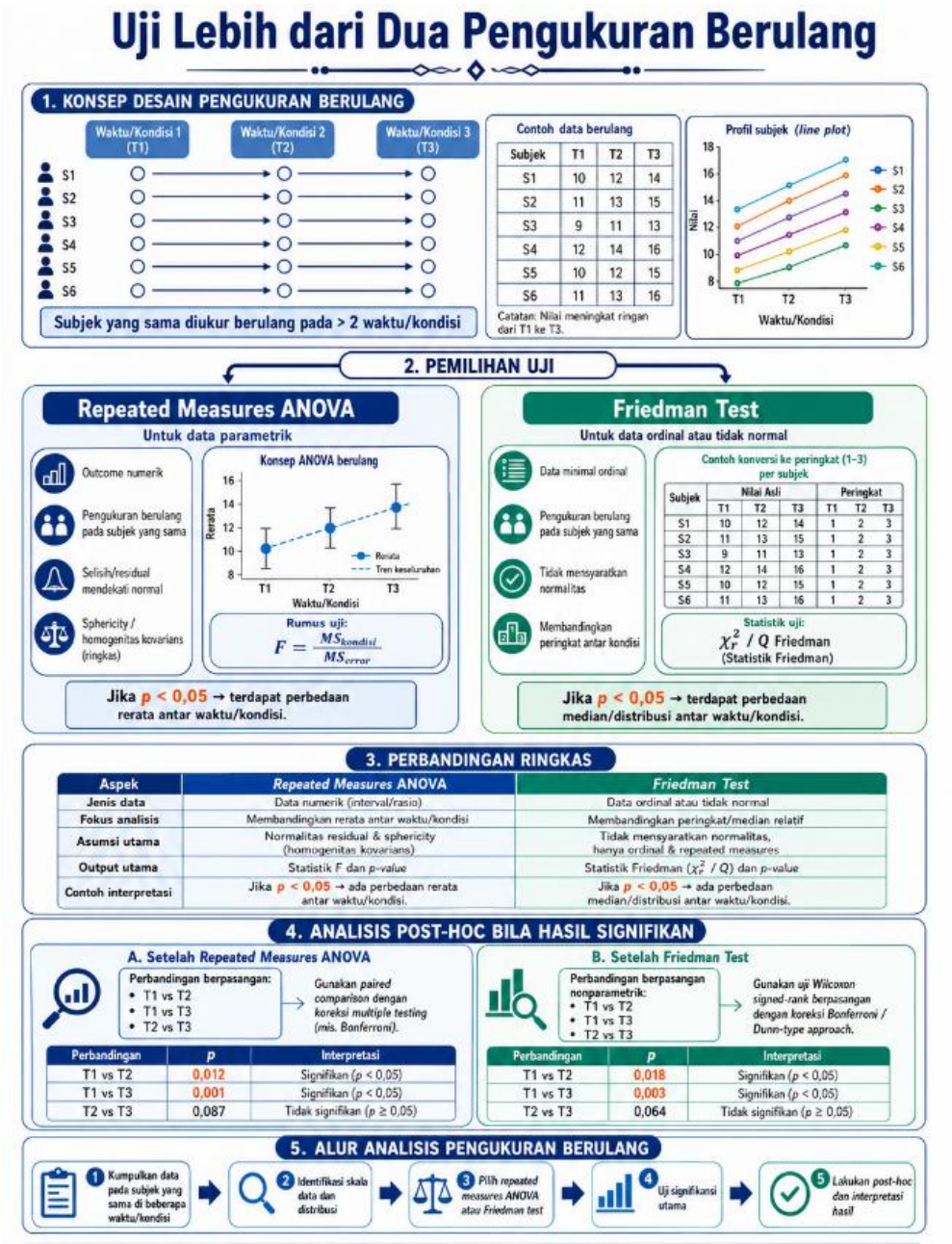
Gambar 5. 8. Kruskal–Wallis test untuk lebih dari dua kelompok

Diagram menunjukkan lebih dari dua kelompok independen, penggabungan seluruh data, pemberian peringkat, perbandingan jumlah peringkat, dan post-hoc nonparametrik. **Sumber:** Diadaptasi dari Chicco *et al*. (2025), Lee (2022), dan Rovetta *et al*. (2025).

5.9 Repeated Measures ANOVA dan Friedman Test

Repeated measures ANOVA digunakan untuk membandingkan rerata outcome numerik pada lebih dari dua pengukuran berulang pada subjek yang sama (Lee, 2022). Contohnya adalah membandingkan tekanan darah pada waktu baseline, 1 bulan, 3 bulan, dan 6 bulan pada pasien yang sama (Lee, 2022). Uji ini mempertimbangkan korelasi antar pengukuran dalam subjek yang sama, sehingga berbeda dari one-way ANOVA biasa yang mengasumsikan kelompok independen (Lee, 2022). Asumsi repeated measures ANOVA mencakup normalitas residual, independensi antar subjek, dan sphericity pada pengukuran berulang (Lee, 2022).

Friedman test adalah alternatif nonparametrik untuk lebih dari dua pengukuran berulang ketika outcome bersifat ordinal atau asumsi repeated measures ANOVA tidak terpenuhi (Chicco *et al*., 2025). Uji ini menggunakan peringkat dalam setiap subjek untuk membandingkan perubahan antar waktu atau kondisi (Chicco *et al*., 2025). Jika Friedman test signifikan, analisis post-hoc berpasangan nonparametrik dapat dilakukan dengan koreksi multiple comparisons (Lee, 2022). Pelaporan uji pengukuran berulang sebaiknya menjelaskan waktu pengukuran, jumlah subjek yang lengkap, missing data, ukuran efek, dan interpretasi klinis perubahan (Rovetta *et al*., 2025).



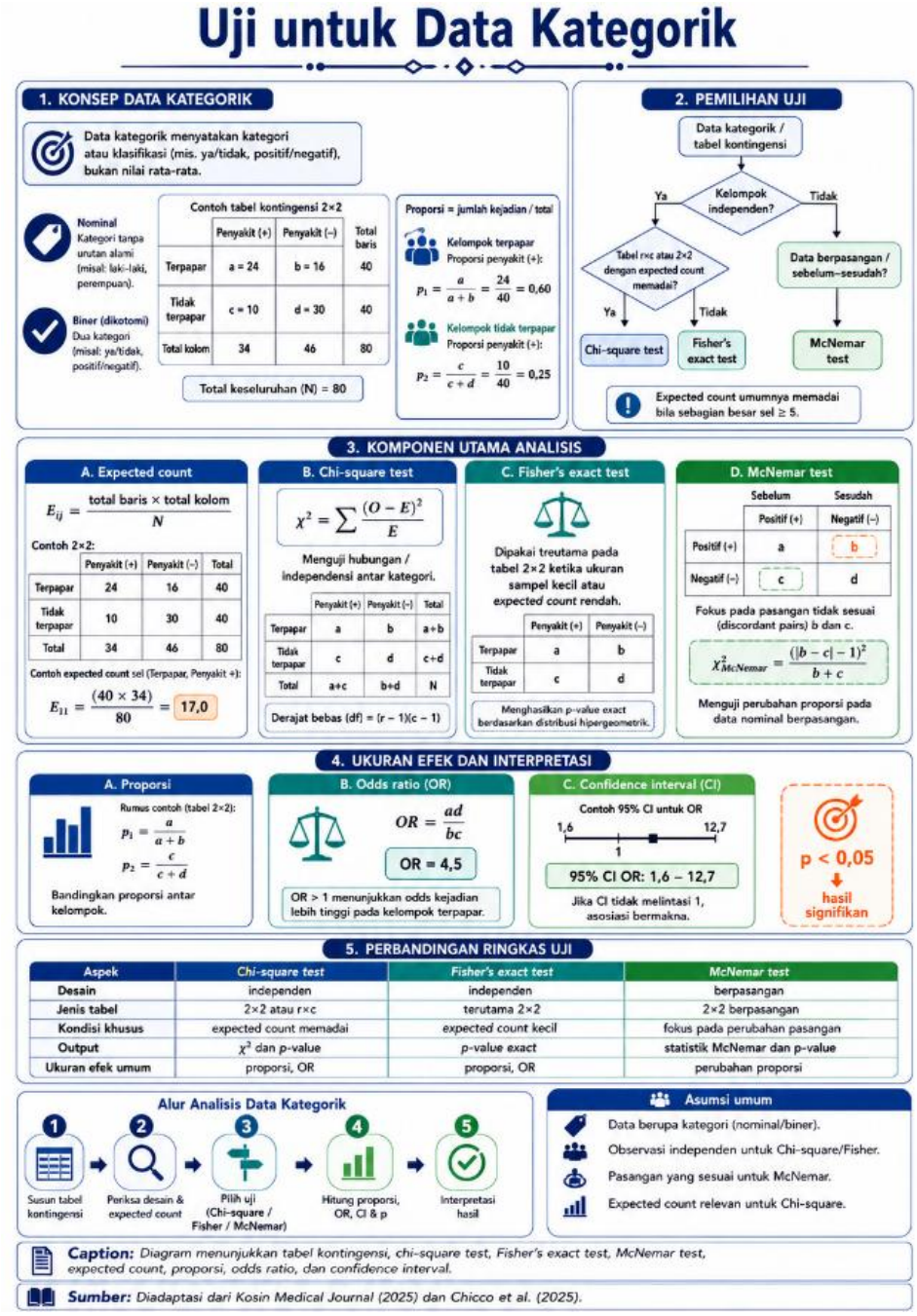
Gambar 5. 9. Uji lebih dari dua pengukuran berulang

Gambar menunjukkan pengukuran berulang pada subjek yang sama, repeated measures ANOVA untuk data parametrik, Friedman test untuk data ordinal atau tidak normal, dan analisis post-hoc. **Sumber:** Diadaptasi dari Lee (2022), Chicco *et al*. (2025), dan Rovetta *et al*. (2025).

5.10 Uji Chi-Square dan Fisher's Exact untuk Data Kategorik

Chi-square test digunakan untuk mengevaluasi hubungan atau perbedaan proporsi antara dua variabel kategorik (Kosin Medical Journal, 2025). Contoh penggunaannya adalah membandingkan proporsi kejadian efek samping antara kelompok terapi dan kelompok kontrol (Kosin Medical Journal, 2025). Uji ini tidak bergantung pada asumsi normalitas outcome numerik karena digunakan pada data frekuensi kategorik (Kosin Medical Journal, 2025). Syarat penting chi-square test adalah expected count dalam sel tabel kontingensi cukup memadai (Chicco *et al*., 2025).

Fisher's exact test digunakan ketika ukuran sampel kecil atau expected count rendah, terutama pada tabel 2×2 (Chicco *et al*., 2025). Fisher's exact test menghitung probabilitas secara eksak berdasarkan konfigurasi tabel, sehingga lebih sesuai ketika asumsi chi-square tidak terpenuhi (Chicco *et al*., 2025). Untuk data kategorik berpasangan, McNemar test dapat digunakan pada tabel 2×2 berpasangan, seperti perubahan status positif-negatif sebelum dan sesudah intervensi (Kosin Medical Journal, 2025). Dengan demikian, data kategorik memiliki jalur analisis tersendiri yang berbeda dari uji parametrik dan nonparametrik untuk data numerik (Kosin Medical Journal, 2025).



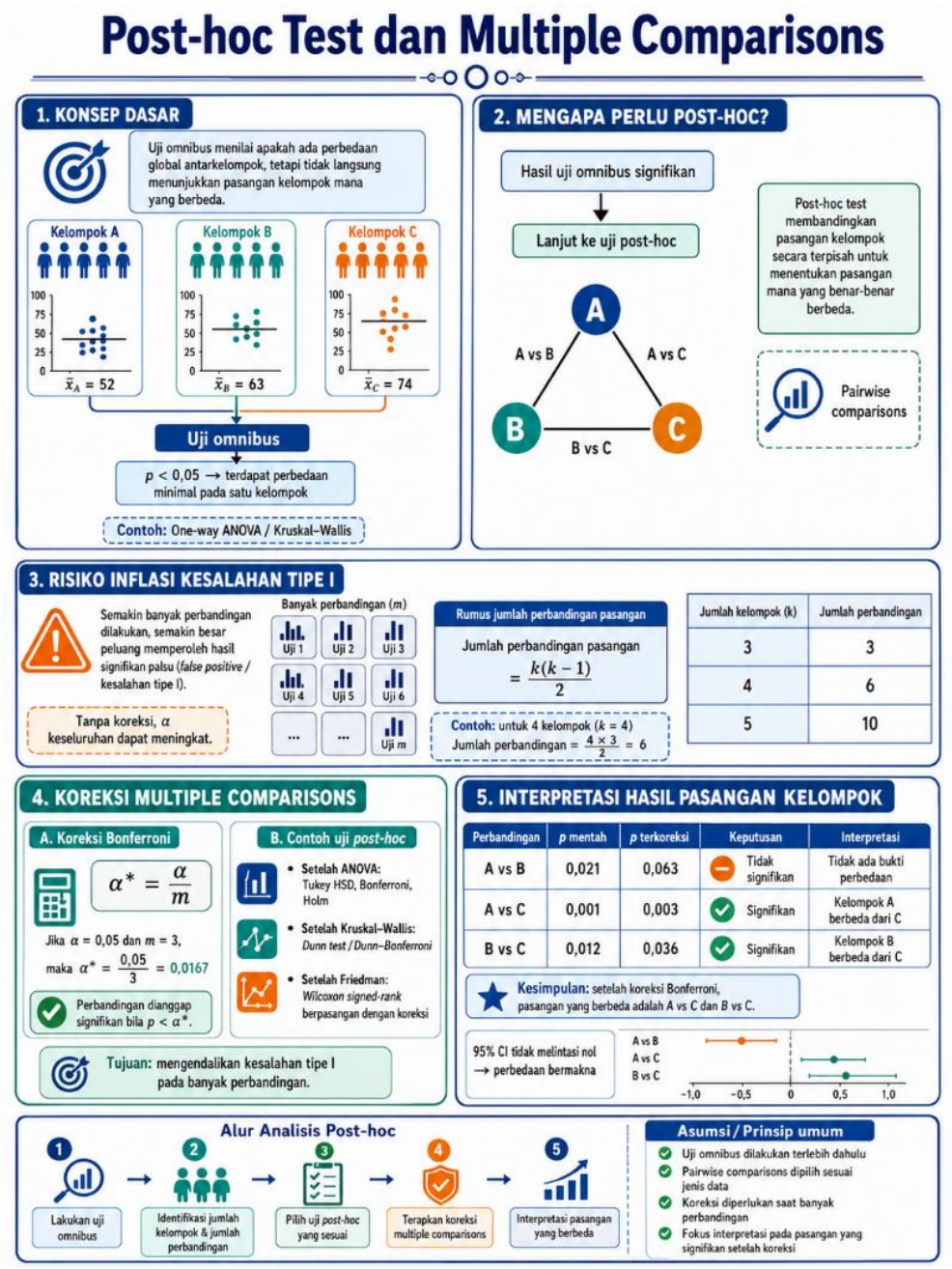
Gambar 5. 10. Uji untuk data kategorik

Diagram menunjukkan tabel kontingensi, chi-square test, Fisher's exact test, McNemar test, expected count, proporsi, odds ratio, dan confidence interval. **Sumber:** Diadaptasi dari Kosin Medical Journal (2025) dan Chicco *et al*. (2025).

5.11 Post-Hoc Test dan Multiple Comparisons

Post-hoc test dilakukan setelah uji omnibus seperti ANOVA atau Kruskal–Wallis menunjukkan adanya perbedaan secara keseluruhan (Lee, 2022). Uji omnibus hanya menunjukkan bahwa paling sedikit ada satu kelompok yang berbeda, tetapi tidak menunjukkan pasangan kelompok mana yang berbeda (Lee, 2022). Oleh karena itu, post-hoc test diperlukan untuk membandingkan pasangan kelompok secara lebih spesifik (Lee, 2022). Contoh post-hoc parametrik meliputi Tukey test, Bonferroni correction, dan Games–Howell test sesuai asumsi varians dan desain penelitian (Lee, 2022).

Multiple comparisons meningkatkan risiko kesalahan tipe I karena semakin banyak perbandingan dilakukan, semakin besar peluang menemukan hasil signifikan secara kebetulan (Lee, 2022). Koreksi seperti Bonferroni atau metode lain dapat digunakan untuk mengendalikan risiko tersebut (Lee, 2022). Pada analisis nonparametrik, post-hoc dapat dilakukan dengan uji pasangan berbasis peringkat disertai koreksi multiple comparisons (Chicco *et al*., 2025). Pelaporan post-hoc sebaiknya mencantumkan metode koreksi, pasangan kelompok yang dibandingkan, nilai p yang telah dikoreksi, confidence interval bila tersedia, dan ukuran efek (Rovetta *et al*., 2025).



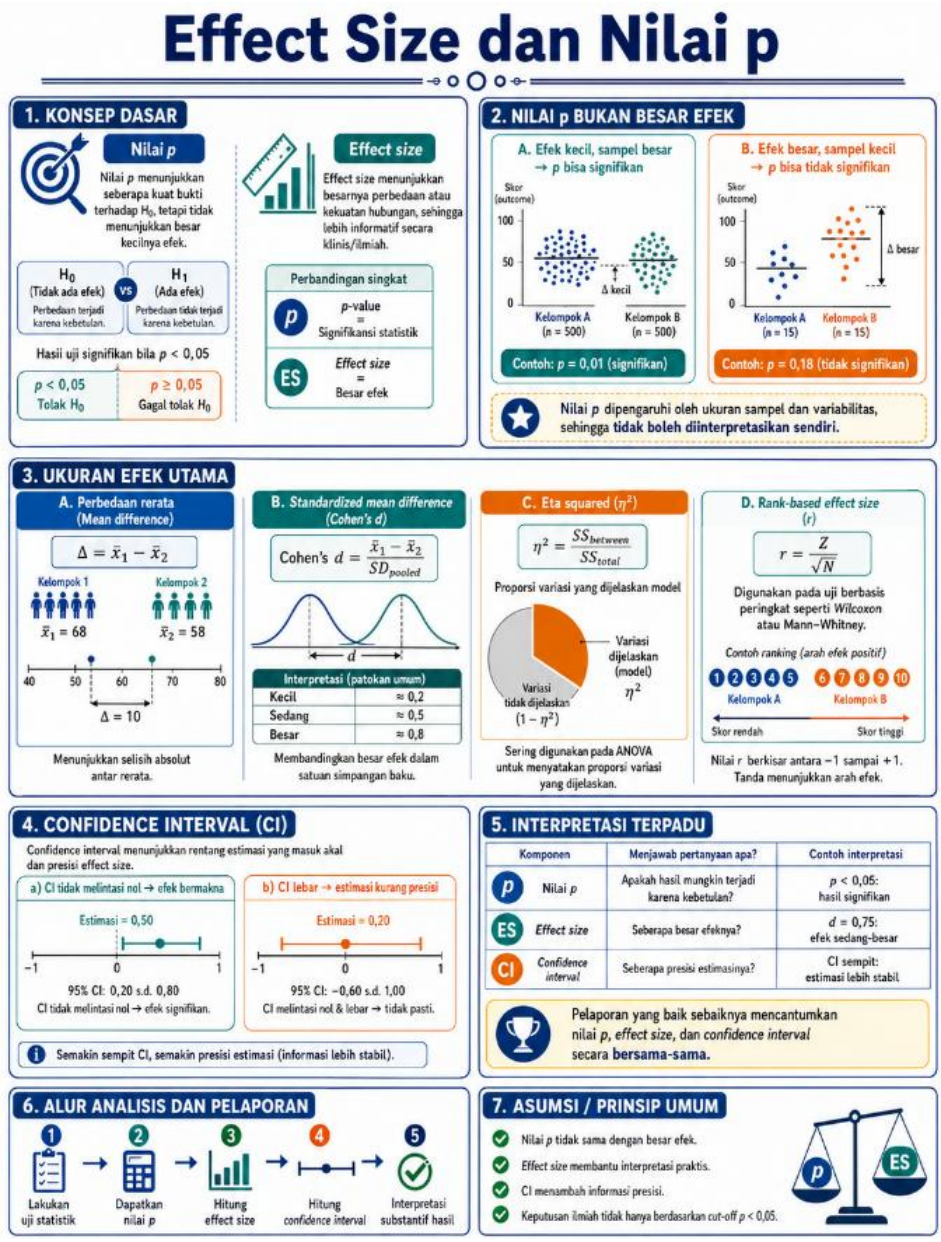
Gambar 5. 11. Post-hoc test dan multiple comparisons

Gambar menunjukkan uji omnibus, perbandingan pasangan kelompok, risiko inflasi kesalahan tipe I, koreksi Bonferroni, dan interpretasi pasangan kelompok yang berbeda. **Sumber:** Diadaptasi dari Lee (2022), Chicco *et al*. (2025), dan Rovetta *et al*. (2025).

5.12 Effect Size dalam Uji Perbedaan Kelompok

Effect size menunjukkan besar perbedaan antar kelompok atau besar perubahan antar pengukuran, sehingga membantu menilai makna praktis hasil statistik (Rovetta *et al*., 2025). Nilai p dapat menunjukkan apakah data kurang kompatibel dengan hipotesis nol, tetapi tidak menunjukkan seberapa besar perbedaan yang terjadi (Rovetta *et al*., 2025). Dalam independent t-test atau paired t-test, effect size dapat dilaporkan sebagai Cohen's d atau standardized mean difference (Rovetta *et al*., 2025). Dalam ANOVA, effect size dapat dilaporkan sebagai eta squared atau partial eta squared (Rovetta *et al*., 2025).

Pada uji nonparametrik, ukuran efek dapat dilaporkan dengan r, rank-biserial correlation, epsilon squared, atau ukuran efek berbasis peringkat lain sesuai uji yang digunakan (Chicco *et al*., 2025). Effect size perlu ditafsirkan bersama konteks klinis atau substantif, karena effect size kecil dapat bermakna bila outcome penting secara klinis, sedangkan effect size besar dapat kurang bermakna bila outcome tidak relevan (Rovetta *et al*., 2025). Confidence interval effect size membantu menunjukkan ketidakpastian estimasi efek (Rovetta *et al*., 2025). Oleh karena itu, pelaporan effect size dan confidence interval meningkatkan kualitas interpretasi dibanding hanya melaporkan "signifikan" atau "tidak signifikan" (Rovetta *et al*., 2025).



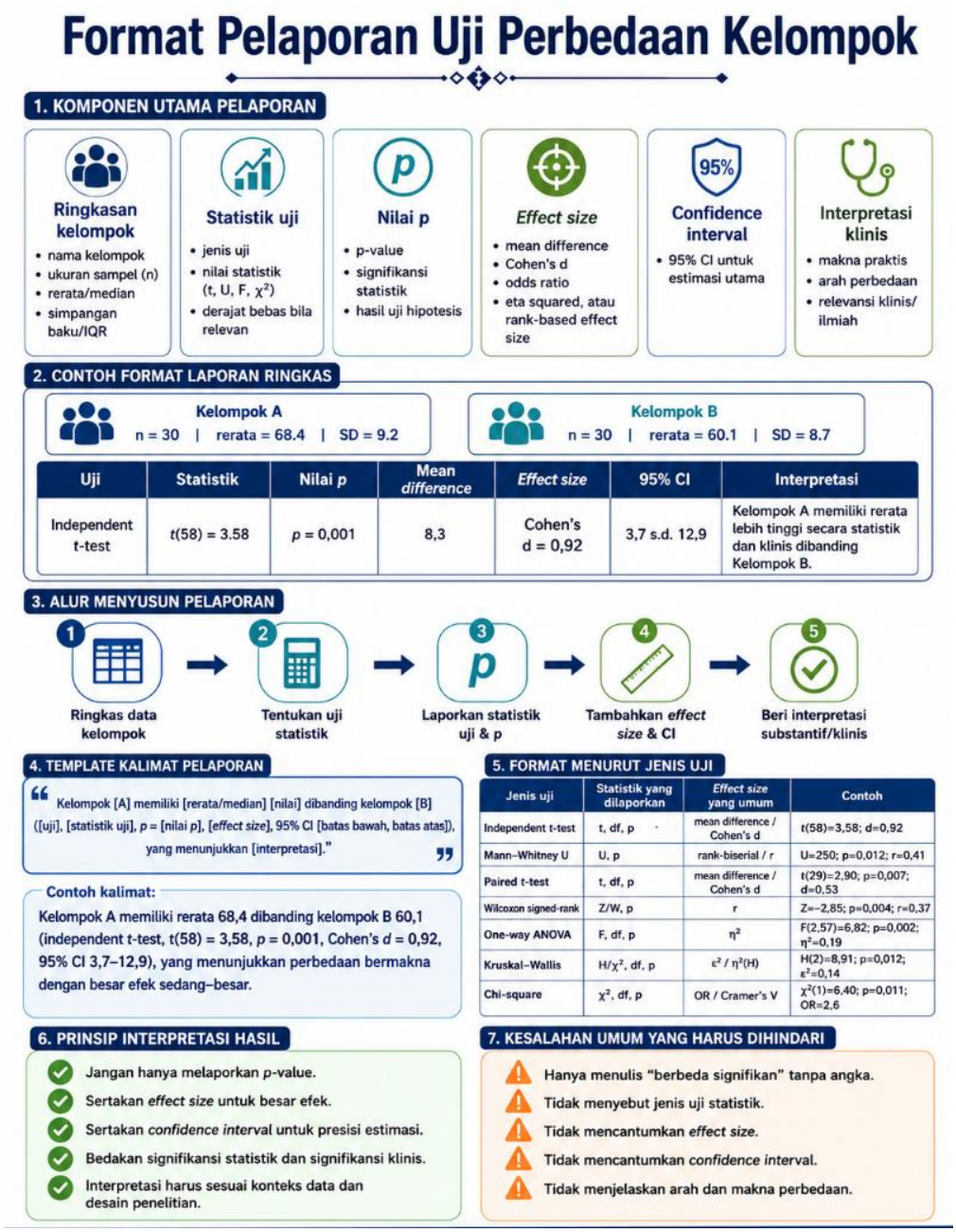
Gambar 5. 12. Effect size dan nilai p

Diagram menunjukkan perbedaan nilai p dan effect size, perbedaan rerata, standardized mean difference, eta squared, rank-based effect size, dan confidence interval. **Sumber:** Diadaptasi dari Rovetta *et al*. (2025) dan Chicco *et al*. (2025).

5.13 Pelaporan Hasil Uji Perbedaan Kelompok

Pelaporan uji perbedaan kelompok harus menyajikan ringkasan data setiap kelompok, nama uji yang digunakan, alasan pemilihan uji, statistik uji, derajat kebebasan bila relevan, nilai p, ukuran efek, dan confidence interval (Rovetta *et al*., 2025). Untuk uji parametrik, data biasanya dilaporkan sebagai mean dan simpangan baku bila distribusi sesuai (Ghasemi & Zahediasl, 2012). Untuk uji nonparametrik, data sering dilaporkan sebagai median dan interquartile range (Cooksey, 2020). Untuk data kategorik, hasil sebaiknya dilaporkan sebagai frekuensi dan persentase disertai ukuran asosiasi seperti risk difference, risk ratio, atau odds ratio bila sesuai (Kosin Medical Journal, 2025).

Kalimat pelaporan yang baik tidak hanya menyebut "terdapat perbedaan signifikan", tetapi menjelaskan arah dan besar perbedaan. Misalnya, "Kelompok intervensi memiliki rerata tekanan darah sistolik lebih rendah dibanding kontrol sebesar 8 mmHg, 95% CI 3–13 mmHg, p = 0,002." Pelaporan seperti ini lebih informatif karena memadukan arah efek, besar efek, ketidakpastian, dan signifikansi statistik (Rovetta *et al*., 2025). Peneliti juga perlu menghindari menyimpulkan tidak ada efek hanya karena p > 0,05 jika confidence interval masih lebar dan ukuran sampel kecil (Rovetta *et al*., 2025).



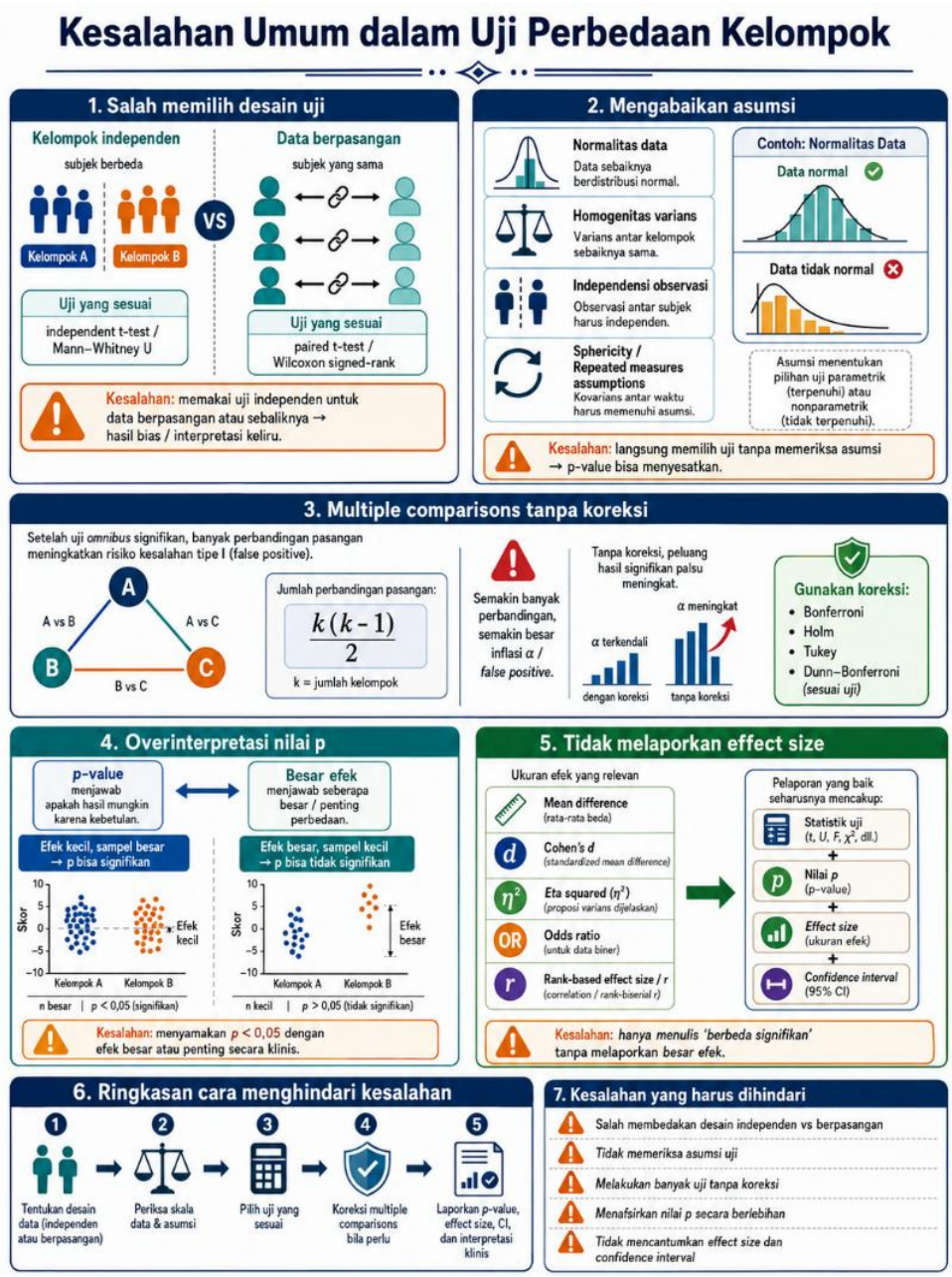
Gambar 5. 13. Format pelaporan uji perbedaan kelompok

Gambar menunjukkan ringkasan kelompok, statistik uji, nilai p, effect size, confidence interval, dan interpretasi klinis hasil. **Sumber:** Diadaptasi dari Rovetta *et al*. (2025), Ghasemi dan Zahediasl (2012), Cooksey (2020), dan Kosin Medical Journal (2025).

5.14 Kesalahan Umum dalam Uji Perbedaan Kelompok

Kesalahan umum pertama adalah menggunakan independent t-test untuk data berpasangan atau paired t-test untuk dua kelompok independen (Chicco *et al*., 2025). Kesalahan umum kedua adalah menggunakan ANOVA untuk data ordinal atau data yang sangat menceng tanpa memeriksa asumsi dan alternatif analisis (Lee, 2022). Kesalahan umum ketiga adalah melakukan banyak uji t antar kelompok setelah ANOVA tanpa koreksi multiple comparisons (Lee, 2022). Kesalahan umum keempat adalah menyimpulkan perbedaan klinis hanya berdasarkan $p < 0,05$ tanpa melihat effect size dan confidence interval (Rovetta *et al*., 2025).

Kesalahan lain adalah memilih uji nonparametrik hanya karena uji normalitas signifikan pada sampel besar, padahal penyimpangan kecil dapat menjadi signifikan secara statistik tetapi tidak penting secara praktis (Ghasemi & Zahediasl, 2012). Peneliti juga sering tidak melaporkan asumsi yang diperiksa sebelum memilih uji parametrik atau nonparametrik (Rovetta *et al*., 2025). Pada data kategorik, kesalahan umum adalah menggunakan chi-square test ketika expected count terlalu rendah, padahal Fisher's exact test lebih sesuai (Chicco *et al*., 2025). Oleh karena itu, pemilihan uji perbedaan kelompok harus dilakukan secara konseptual, bukan mekanis (Chicco *et al*., 2025).

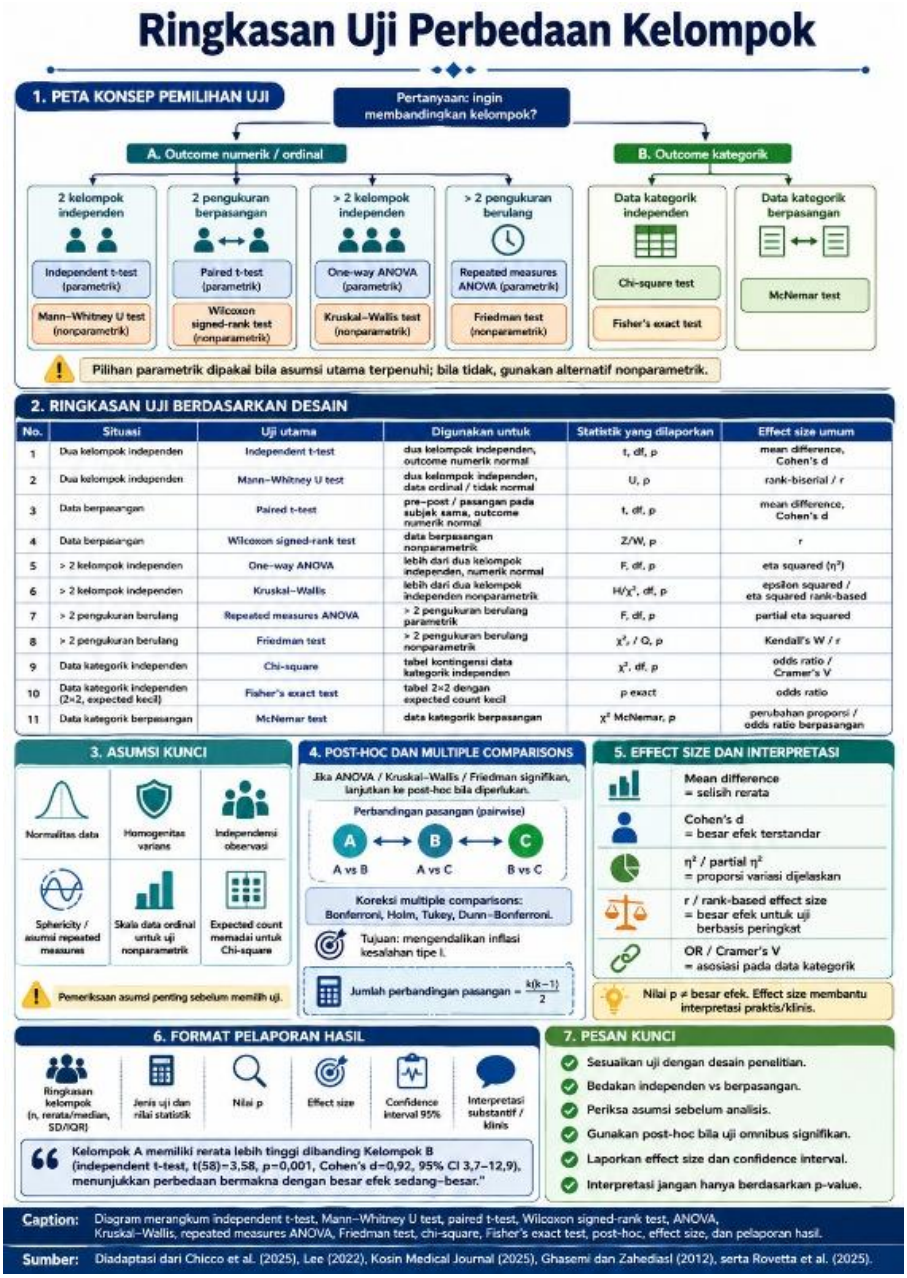


Gambar 5. 14. Kesalahan umum dalam uji perbedaan kelompok

Diagram menunjukkan kesalahan memilih uji independen versus berpasangan, mengabaikan asumsi, multiple comparisons tanpa koreksi, overinterpretasi nilai p, dan tidak melaporkan effect size. **Sumber:** Diadaptasi dari Chicco *et al*. (2025), Lee (2022), Ghasemi dan Zahediasl (2012), serta Rovetta *et al*. (2025).

Bab ini menjelaskan bahwa uji perbedaan kelompok digunakan untuk membandingkan outcome antara dua kelompok atau lebih sesuai desain penelitian (Lee, 2022). Independent t-test dan Mann–Whitney U test digunakan untuk dua kelompok independen, sedangkan paired t-test dan Wilcoxon signed-rank test digunakan untuk dua pengukuran berpasangan (Chicco *et al*., 2025). One-way ANOVA dan Kruskal–Wallis test digunakan untuk lebih dari dua kelompok independen sesuai asumsi parametrik atau nonparametrik (Chicco *et al*., 2025). Repeated measures ANOVA dan Friedman test digunakan untuk lebih dari dua pengukuran berulang pada subjek yang sama (Lee, 2022; Chicco *et al*., 2025).

Bab ini juga menegaskan bahwa pemilihan uji tidak boleh hanya berdasarkan nilai p uji normalitas (Ghasemi & Zahediasl, 2012). Peneliti perlu mempertimbangkan jenis outcome, jumlah kelompok, hubungan antarobservasi, distribusi data, outlier, homogenitas varians, dan tujuan interpretasi (Chicco *et al*., 2025). Pelaporan hasil sebaiknya mencakup ringkasan data, statistik uji, nilai p, effect size, confidence interval, dan makna klinis atau substantif (Rovetta *et al*., 2025). Dengan demikian, uji perbedaan kelompok bukan sekadar prosedur teknis, tetapi bagian dari penalaran ilmiah yang harus disesuaikan dengan desain penelitian dan karakteristik data (Rovetta *et al*., 2025).



Gambar 5. 15. Ringkasan uji perbedaan kelompok

Diagram merangkum independent t-test, Mann–Whitney U test, paired t-test, Wilcoxon signed-rank test, ANOVA, Kruskal–Wallis, repeated measures ANOVA, Friedman test, chi-square, Fisher's exact test, post-hoc, effect size, dan pelaporan hasil. **Sumber:** Diadaptasi dari Chicco *et al*. (2025), Lee (2022), Kosin Medical Journal (2025), Ghasemi dan Zahediasl (2012), serta Rovetta *et al*. (2025).

Review Questions

1. **Apa tujuan utama uji perbedaan kelompok?**
   Jawaban: Tujuannya adalah menilai apakah outcome berbeda antara dua kelompok atau lebih sesuai desain penelitian.
2. **Kapan independent t-test digunakan?**
   Jawaban: Independent t-test digunakan untuk membandingkan rerata outcome numerik antara dua kelompok independen yang memenuhi asumsi parametrik.
3. **Apa alternatif nonparametrik untuk independent t-test?**
   Jawaban: Alternatifnya adalah Mann–Whitney U test.
4. **Kapan paired t-test digunakan?**
   Jawaban: Paired t-test digunakan untuk membandingkan rerata dua pengukuran berpasangan pada subjek yang sama.
5. **Apa alternatif nonparametrik untuk paired t-test?**
   Jawaban: Alternatifnya adalah Wilcoxon signed-rank test.
6. **Kapan one-way ANOVA digunakan?**
   Jawaban: One-way ANOVA digunakan untuk membandingkan rerata outcome numerik pada lebih dari dua kelompok independen yang memenuhi asumsi parametrik.
7. **Apa alternatif nonparametrik untuk one-way ANOVA?**
   Jawaban: Alternatifnya adalah Kruskal–Wallis test.
8. **Apa fungsi post-hoc test?**
   Jawaban: Post-hoc test digunakan untuk mengetahui pasangan kelompok mana yang berbeda setelah uji omnibus menunjukkan hasil signifikan.
9. **Mengapa effect size penting?**
   Jawaban: Effect size penting karena menunjukkan besar perbedaan atau efek, sedangkan nilai p hanya menunjukkan kompatibilitas data dengan hipotesis nol.
10. **Apa kesalahan umum dalam uji perbedaan kelompok?**
    Jawaban: Kesalahan umum meliputi salah membedakan desain independen dan berpasangan, mengabaikan asumsi, melakukan banyak perbandingan tanpa koreksi, serta menafsirkan hasil hanya berdasarkan nilai p.

DAFTAR PUSTAKA

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*, Article 40. https://doi.org/10.1186/s13040-025-00465-6

Cooksey, R. W. (2020). Descriptive statistics for summarising data. In *Illustrating statistical procedures: Finding meaning in quantitative data* (pp. 61–139). Springer.

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

Kosin Medical Journal. (2025). Common statistical methods used in medical research. *Kosin Medical Journal*.

Lee, S. W. (2022). Methods for testing statistical differences between groups in medical research. *Life Cycle, 2*, e1.

Rovetta, A., Piretta, L., & Mansournia, M. A. (2025). p-values and confidence intervals as compatibility measures: Guidelines for interpreting statistical studies in clinical research. *The Lancet Regional Health – Southeast Asia*.

BAB 6
ANALISIS KORELASI DAN REGRESI: PEARSON, SPEARMAN, REGRESI LINEAR, REGRESI LOGISTIK, DAN INTERPRETASI MODEL STATISTIK

6.1 Pendahuluan

Analisis korelasi dan regresi digunakan ketika peneliti ingin memahami hubungan antarvariabel, baik dalam bentuk asosiasi sederhana maupun model prediksi yang lebih kompleks. Korelasi digunakan untuk menilai arah dan kekuatan hubungan antara dua variabel, sedangkan regresi digunakan untuk menjelaskan atau memprediksi nilai outcome berdasarkan satu atau lebih variabel prediktor (Wisniewski, 2024). Dalam penelitian kesehatan, korelasi dapat digunakan untuk menilai hubungan antara kadar biomarker dan tekanan darah, sedangkan regresi dapat digunakan untuk memprediksi risiko penyakit berdasarkan usia, indeks massa tubuh, status merokok, atau variabel klinis lain. Namun, korelasi dan regresi tidak otomatis membuktikan hubungan sebab-akibat karena kausalitas membutuhkan desain penelitian, kontrol confounding, temporality, dan argumentasi biologis yang kuat (Ahmed, 2025).

Bab ini membahas korelasi Pearson, korelasi Spearman, korelasi Kendall, regresi linear sederhana, regresi linear berganda, regresi logistik, odds ratio, asumsi model, interpretasi koefisien, evaluasi model, confounding, multikolinearitas, overfitting, dan pelaporan hasil. Pemilihan metode korelasi atau regresi harus disesuaikan dengan jenis outcome, bentuk hubungan, distribusi data, skala pengukuran, dan tujuan analisis. Pearson lebih sesuai untuk hubungan linear antara dua variabel numerik, sedangkan Spearman lebih sesuai untuk data ordinal atau hubungan monotonic yang tidak harus linear (Wisniewski, 2024; Bujang, 2024).



Gambar 6. 1. Peta konsep korelasi dan regresi dalam penelitian
Kesehatan

Ilustrasi menunjukkan korelasi Pearson, Spearman, Kendall, regresi linear, regresi logistik, variabel prediktor, outcome, koefisien, confidence interval, nilai p, dan interpretasi klinis. **Sumber:** Diadaptasi dari Wisniewski (2024), Chicco *et al*. (2025), Bujang (2024), Zabor *et al*. (2022), dan Collins *et al*. (2024).

6.2 Konsep Dasar Asosiasi Antarvariabel

Asosiasi antarvariabel menunjukkan bahwa perubahan nilai satu variabel berkaitan dengan perubahan nilai variabel lain. Asosiasi dapat bersifat positif, negatif, atau tidak jelas. Asosiasi positif berarti nilai satu variabel cenderung meningkat ketika variabel lain meningkat. Asosiasi negatif berarti nilai satu variabel cenderung menurun ketika variabel lain meningkat. Korelasi menggambarkan arah dan kekuatan asosiasi, tetapi tidak menjelaskan apakah satu variabel menyebabkan perubahan variabel lain (Ahmed, 2025).

Dalam penelitian klinis, hubungan antara dua variabel sering dipengaruhi oleh faktor ketiga. Misalnya, hubungan antara aktivitas fisik dan tekanan darah dapat dipengaruhi oleh usia, berat badan, konsumsi obat, atau pola makan. Faktor seperti ini disebut confounder apabila berhubungan dengan paparan dan outcome serta mengubah estimasi

hubungan utama. Oleh karena itu, analisis asosiasi sederhana sering perlu dilanjutkan dengan regresi multivariat untuk mengontrol confounding dan memperkirakan hubungan yang lebih teradjustment. Model prediksi klinis, termasuk model berbasis regresi dan machine learning, perlu dilaporkan secara transparan sesuai pedoman TRIPOD+AI (Collins *et al*., 2024).



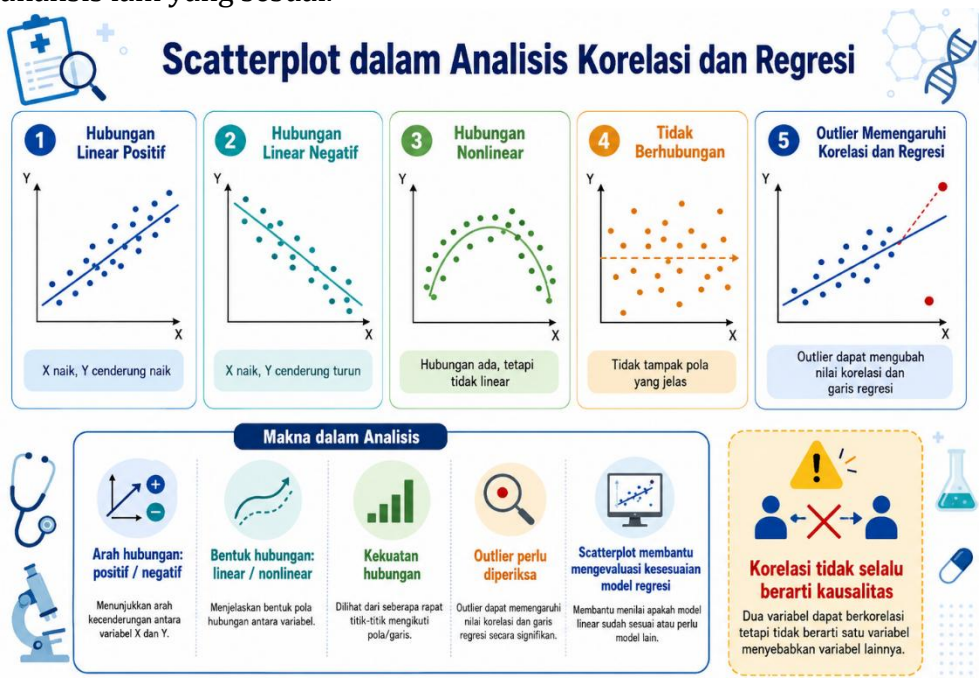
Gambar 6. 2. Asosiasi, korelasi, dan kausalitas

Diagram menunjukkan hubungan antarvariabel, korelasi positif, korelasi negatif, confounder, regresi teradjustment, dan perbedaan antara asosiasi serta kausalitas. **Sumber:** Diadaptasi dari Ahmed (2025), Wisniewski (2024), Rovetta *et al*. (2025), dan Collins *et al*. (2024).

6.3 Scatterplot sebagai Langkah Awal Analisis Hubungan

Scatterplot adalah grafik yang menampilkan pasangan nilai dari dua variabel numerik pada sumbu X dan Y. Scatterplot membantu peneliti melihat apakah hubungan tampak linear, nonlinear, monotonic, heterogen, atau dipengaruhi oleh outlier. Sebelum menghitung korelasi Pearson atau membuat regresi linear, scatterplot sebaiknya diperiksa karena angka korelasi tunggal dapat menyembunyikan pola data yang kompleks. Hubungan yang tampak kuat pada angka korelasi dapat sebenarnya didorong oleh beberapa outlier ekstrem.

Scatterplot juga membantu mendeteksi heteroskedastisitas, yaitu kondisi ketika variasi outcome berubah pada tingkat prediktor yang berbeda. Dalam regresi linear, pola residual yang melebar atau menyempit dapat menunjukkan pelanggaran asumsi homoskedastisitas. Jika pola hubungan tidak linear, model linear sederhana dapat menyesatkan. Dalam kondisi tersebut, peneliti dapat mempertimbangkan transformasi, model nonlinear, spline, generalized linear model, atau analisis lain yang sesuai.



Gambar 6. 3. Scatterplot dalam analisis korelasi dan regresi

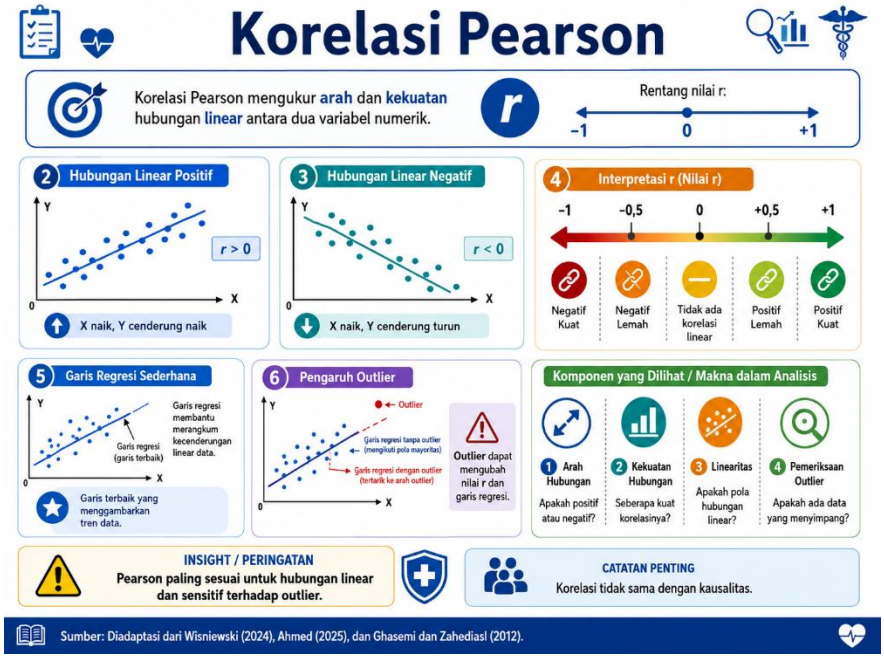
Gambar menunjukkan hubungan linear positif, linear negatif, nonlinear, tidak berhubungan, serta outlier yang dapat memengaruhi korelasi dan garis regresi. **Sumber:** Diadaptasi dari Wisniewski (2024), Ghasemi dan Zahediasl (2012), serta Kosin Medical Journal (2025).

6.4 Korelasi Pearson

Korelasi Pearson digunakan untuk mengukur kekuatan dan arah hubungan linear antara dua variabel numerik. Nilai koefisien Pearson, yang sering ditulis sebagai r, berkisar dari -1 sampai +1. Nilai r mendekati +1 menunjukkan hubungan linear positif yang kuat, nilai r mendekati -1 menunjukkan hubungan linear negatif yang kuat, dan nilai r mendekati 0 menunjukkan tidak adanya hubungan linear yang jelas. Korelasi Pearson cocok digunakan ketika variabel berskala interval atau rasio, hubungan

bersifat linear, dan outlier tidak mendominasi pola data (Wisniewski, 2024).

Korelasi Pearson tidak berarti kedua variabel memiliki hubungan sebab-akibat. Misalnya, korelasi antara kadar kolesterol dan usia menunjukkan bahwa kedua variabel bergerak bersama dalam data, tetapi tidak otomatis membuktikan bahwa usia secara langsung menyebabkan perubahan kolesterol. Interpretasi Pearson harus memperhatikan scatterplot, outlier, rentang data, ukuran sampel, dan konteks biologis. Jika hubungan tidak linear atau data memiliki outlier ekstrem, Pearson dapat memberikan estimasi yang menyesatkan.



Gambar 6. 4. Korelasi Pearson

Diagram menunjukkan hubungan linear positif, linear negatif, nilai r, garis regresi sederhana, dan pengaruh outlier terhadap koefisien Pearson. **Sumber:** Diadaptasi dari Wisniewski (2024), Ahmed (2025), dan Ghasemi dan Zahediasl (2012).

6.5 Korelasi Spearman

Korelasi Spearman digunakan untuk menilai hubungan monotonic antara dua variabel berdasarkan peringkat data. Hubungan monotonic berarti nilai satu variabel cenderung meningkat atau menurun secara konsisten ketika variabel lain meningkat, tetapi pola tersebut tidak harus linear sempurna. Spearman lebih sesuai untuk data ordinal, data numerik yang tidak normal, atau data yang memiliki outlier yang dapat

mengganggu Pearson. Bujang (2024) membahas ukuran sampel untuk Pearson, Spearman, dan Kendall berdasarkan effect size serta lebar confidence interval, sehingga pemilihan korelasi juga perlu dikaitkan dengan tujuan estimasi dan presisi (Bujang, 2024).

Spearman sering digunakan pada skor klinis, tingkat nyeri, skor kualitas hidup, tingkat kepuasan, atau data laboratorium yang sangat menceng. Karena Spearman berbasis peringkat, nilai ekstrem tidak memengaruhi hasil sekuat pada Pearson. Namun, Spearman tetap perlu ditafsirkan dengan scatterplot karena hubungan yang sangat kompleks tidak selalu dapat dirangkum secara baik oleh satu koefisien korelasi. Jika data memiliki banyak nilai yang sama atau ties, interpretasi Spearman perlu lebih hati-hati.



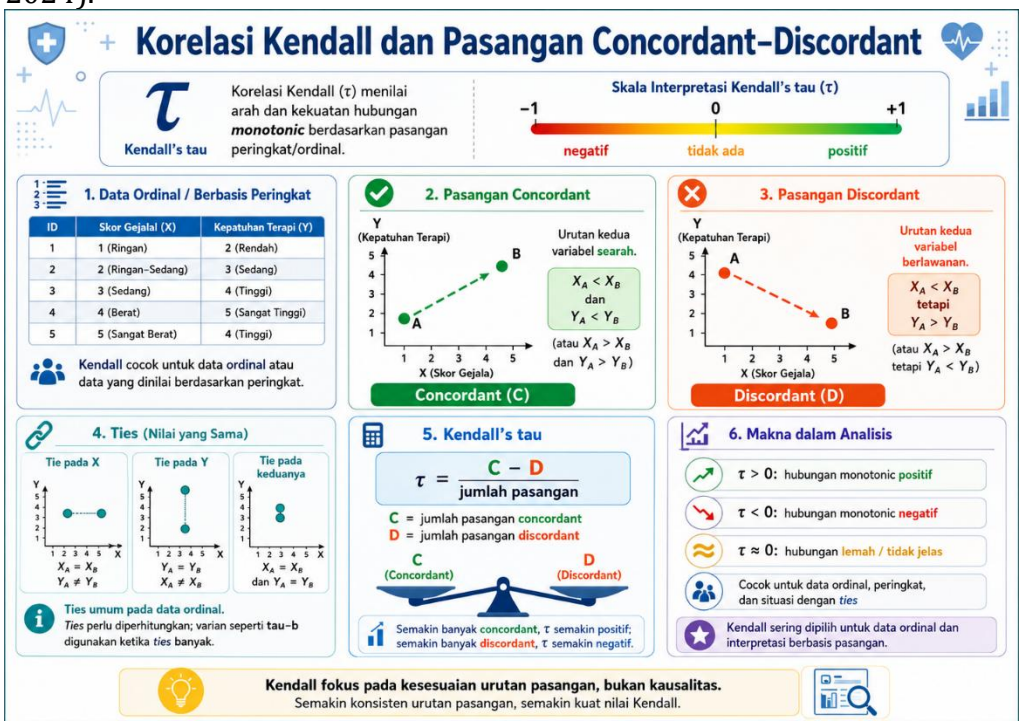
Gambar 6. 5. Korelasi Spearman berbasis peringkat

Gambar menunjukkan perubahan nilai asli menjadi peringkat, hubungan monotonic meningkat, hubungan monotonic menurun, dan perbedaan Spearman dari Pearson. **Sumber:** Diadaptasi dari Wisniewski (2024), Bujang (2024), dan Chicco *et al*. (2025).

6.6 Korelasi Kendall

Korelasi Kendall, terutama Kendall's tau, digunakan untuk menilai asosiasi ordinal berdasarkan pasangan data concordant dan discordant. Pasangan concordant terjadi ketika urutan dua observasi konsisten pada kedua variabel, sedangkan pasangan discordant terjadi ketika urutannya berlawanan. Kendall sering digunakan untuk data ordinal, sampel kecil, atau data dengan banyak ties. Dalam beberapa konteks, Kendall dapat lebih stabil dibanding Spearman ketika ukuran sampel kecil atau banyak nilai yang sama.

Kendall tidak sepopuler Pearson dan Spearman dalam laporan klinis umum, tetapi berguna ketika peneliti ingin menilai konsistensi urutan antarvariabel. Misalnya, Kendall dapat digunakan untuk menilai hubungan antara ranking tingkat keparahan penyakit dan ranking skor kualitas hidup. Bujang (2024) memasukkan Kendall's tau-b dalam pembahasan ukuran sampel korelasi, yang menunjukkan bahwa metode ini tetap penting dalam perencanaan analisis asosiasi ordinal (Bujang, 2024).



Gambar 6. 6. Korelasi Kendall dan pasangan concordant–discordant
Diagram menunjukkan pasangan concordant, pasangan discordant, ties, Kendall's tau, dan penggunaannya pada data ordinal. **Sumber:** Diadaptasi dari Bujang (2024), Wisniewski (2024), dan prinsip analisis korelasi ordinal.

6.7 Interpretasi Koefisien Korelasi

Koefisien korelasi menunjukkan arah dan kekuatan hubungan, tetapi interpretasi kekuatan hubungan tidak boleh dilakukan secara mekanis. Nilai r sebesar 0,30 dapat dianggap kecil dalam satu konteks, tetapi dapat bermakna dalam penelitian klinis bila outcome penting atau intervensi mudah dilakukan. Sebaliknya, korelasi tinggi belum tentu bermakna klinis bila variabel tidak relevan atau hubungan dipengaruhi oleh confounding. Oleh karena itu, interpretasi korelasi harus menggabungkan nilai koefisien, confidence interval, nilai p, scatterplot, dan konteks penelitian.

Nilai p pada korelasi hanya menunjukkan kompatibilitas data dengan hipotesis nol bahwa korelasi populasi sama dengan nol. Nilai p tidak menunjukkan besar hubungan, makna klinis, atau bukti kausalitas. Confidence interval korelasi lebih informatif karena menunjukkan rentang nilai korelasi yang masuk akal berdasarkan data sampel. Pelaporan korelasi yang baik sebaiknya menyertakan jenis korelasi, koefisien, confidence interval, nilai p, jumlah sampel, dan alasan pemilihan metode korelasi.

Tabel 6. 1. Interpretasi umum korelasi

Nilai absolut korelasi	Interpretasi umum	Catatan ilmiah
0,00–0,19	Sangat lemah	Dapat tetap penting bila outcome klinis sangat relevan
0,20–0,39	Lemah	Perlu konteks dan confidence interval
0,40–0,59	Sedang	Perlu evaluasi scatterplot dan outlier
0,60–0,79	Kuat	Tidak otomatis berarti kausal
0,80–1,00	Sangat kuat	Perlu waspada duplikasi variabel atau pengukuran serupa

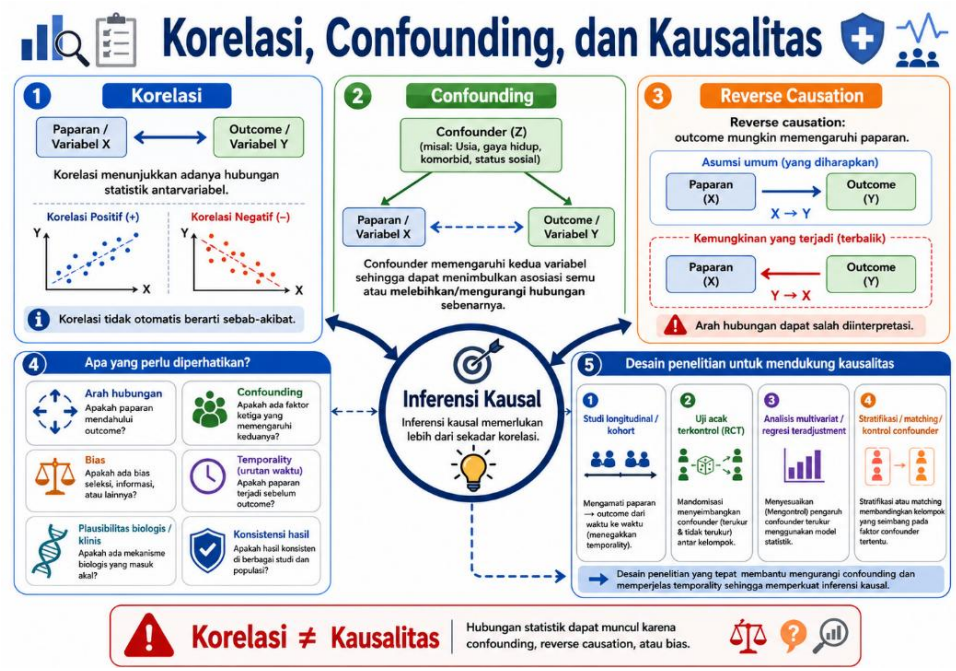
Sumber: Diadaptasi dari Wisniewski (2024), Ahmed (2025), dan Rovetta *et al*. (2025).

6.8 Korelasi Tidak Sama dengan Kausalitas

Korelasi tidak membuktikan kausalitas karena dua variabel dapat berkorelasi akibat confounding, reverse causation, bias seleksi, atau kebetulan statistik. Misalnya, hubungan antara konsumsi kopi dan tekanan darah dapat dipengaruhi oleh usia, stres, aktivitas fisik, obat

antihipertensi, atau kebiasaan merokok. Dalam penelitian observasional, klaim kausal memerlukan bukti temporality, kekuatan asosiasi, konsistensi temuan, plausibilitas biologis, dose-response, dan kontrol confounding. Ahmed (2025) menekankan bahwa korelasi harus dibedakan dari kausalitas agar hasil statistik tidak disalahartikan dalam praktik klinis (Ahmed, 2025).

Regresi dapat membantu mengontrol beberapa confounder, tetapi regresi tidak otomatis menyelesaikan semua masalah kausalitas. Confounder yang tidak diukur tetap dapat menyebabkan bias. Kesalahan pengukuran, selection bias, dan missing data juga dapat mengganggu interpretasi hubungan. Oleh karena itu, analisis statistik harus selalu dihubungkan dengan desain studi, kualitas data, dan teori substantif.



Gambar 6. 7. Korelasi, confounding, dan kausalitas

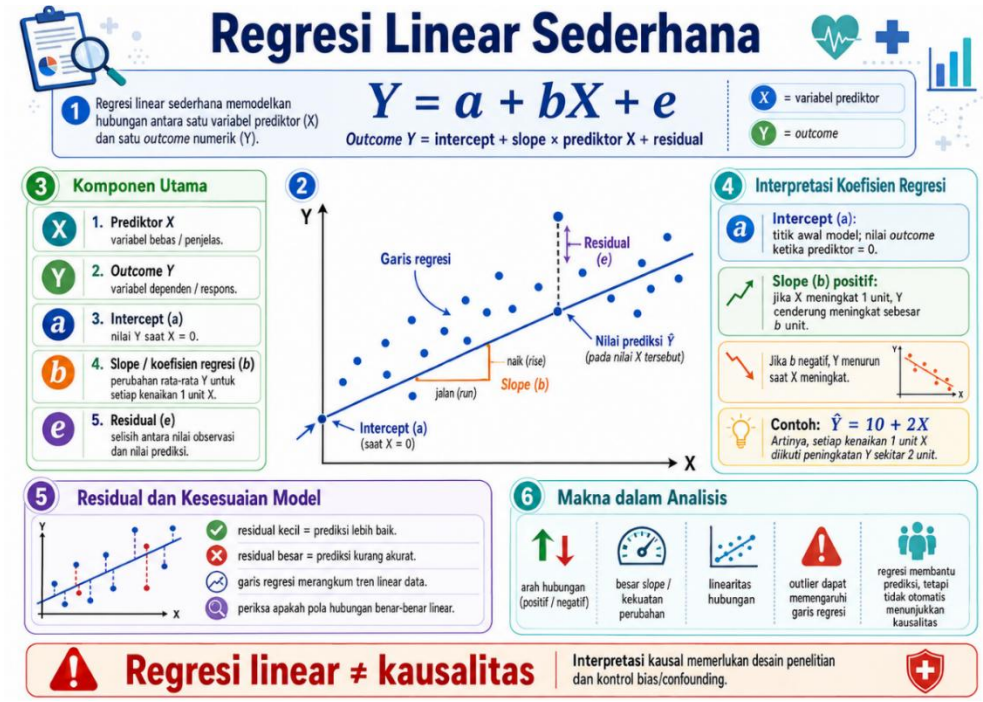
Gambar menunjukkan dua variabel yang berkorelasi, confounder yang memengaruhi keduanya, reverse causation, dan kebutuhan desain penelitian untuk mendukung inferensi kausal. **Sumber:** Diadaptasi dari Ahmed (2025), Rovetta *et al*. (2025), dan prinsip epidemiologi klinis.

6.9 Regresi Linear Sederhana

Regresi linear sederhana digunakan untuk memodelkan hubungan antara satu variabel prediktor dan satu outcome numerik kontinu. Model ini menghasilkan persamaan garis yang memperkirakan nilai outcome

berdasarkan nilai prediktor. Koefisien regresi menunjukkan perubahan rata-rata outcome untuk setiap kenaikan satu unit prediktor. Jika koefisien bernilai positif, outcome cenderung meningkat ketika prediktor meningkat; jika koefisien bernilai negatif, outcome cenderung menurun ketika prediktor meningkat.

Regresi linear sederhana berbeda dari korelasi karena regresi memiliki arah model yang jelas, yaitu prediktor digunakan untuk menjelaskan atau memprediksi outcome. Misalnya, usia dapat digunakan sebagai prediktor tekanan darah, tetapi tekanan darah tidak otomatis menjadi prediktor usia dalam model yang sama. Model regresi linear perlu diperiksa melalui scatterplot, residual plot, Q–Q plot residual, dan evaluasi outlier berpengaruh. Kosin Medical Journal (2025) menjelaskan bahwa metode regresi dalam penelitian medis dipilih berdasarkan jenis outcome, termasuk regresi linear untuk outcome kontinu dan regresi logistik untuk outcome dikotom (Kosin Medical Journal, 2025).



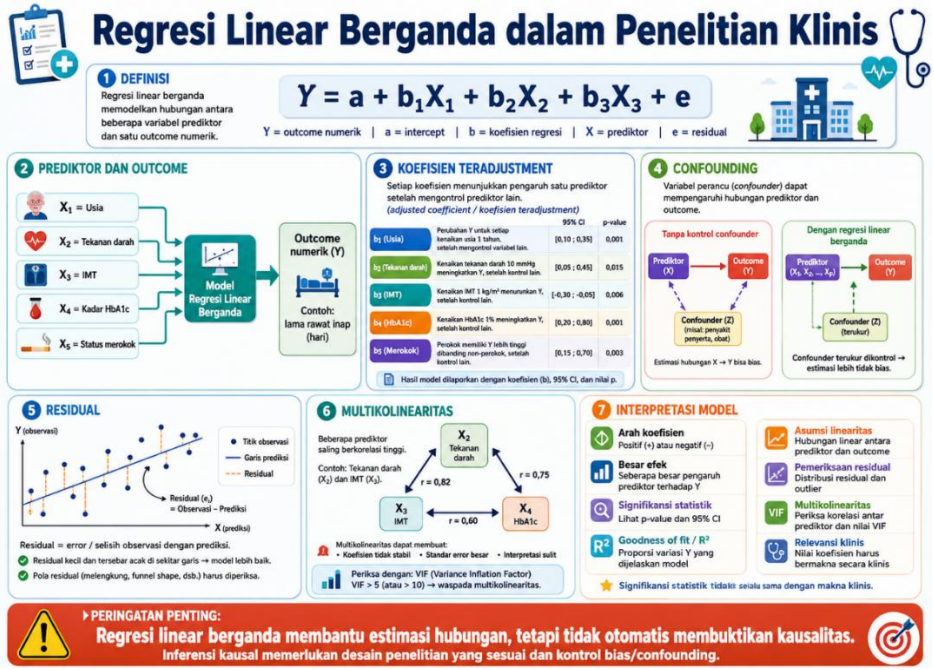
Gambar 6. 8. Regresi linear sederhana

Diagram menunjukkan variabel prediktor X, outcome Y, garis regresi, intercept, slope, residual, dan interpretasi koefisien regresi. **Sumber:** Diadaptasi dari Wisniewski (2024), Kosin Medical Journal (2025), dan Ghasemi dan Zahediasl (2012).

6.10 Regresi Linear Berganda

Regresi linear berganda digunakan ketika peneliti ingin menjelaskan outcome numerik kontinu dengan lebih dari satu prediktor. Model ini memungkinkan peneliti memperkirakan hubungan antara satu prediktor dan outcome setelah mengontrol prediktor lain dalam model. Misalnya, tekanan darah sistolik dapat dimodelkan berdasarkan usia, indeks massa tubuh, status merokok, aktivitas fisik, dan konsumsi obat. Koefisien setiap prediktor menunjukkan perubahan rata-rata outcome untuk setiap kenaikan satu unit prediktor ketika prediktor lain dianggap konstan.

Regresi linear berganda sangat berguna untuk mengontrol confounding, tetapi tetap memerlukan asumsi dan evaluasi model. Asumsi yang perlu diperhatikan meliputi linearitas, independensi residual, homoskedastisitas, normalitas residual, dan tidak adanya multikolinearitas berlebihan. Jika prediktor saling berkorelasi sangat tinggi, koefisien regresi dapat menjadi tidak stabil. Model yang terlalu kompleks untuk ukuran sampel yang kecil juga dapat mengalami overfitting.



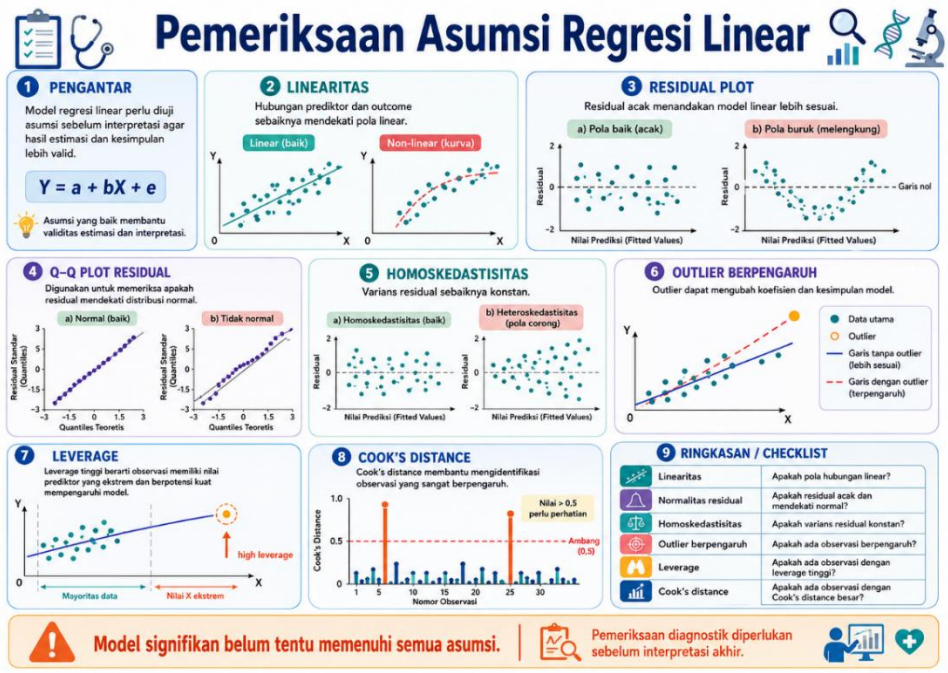
Gambar 6. 9. Regresi linear berganda dalam penelitian klinis

Gambar menunjukkan beberapa prediktor, outcome numerik, koefisien teradjustment, confounding, residual, multikolinearitas, dan interpretasi model. **Sumber:** Diadaptasi dari Kosin Medical Journal (2025), Collins *et al*. (2024), dan Rovetta *et al*. (2025).

6.11 Asumsi Regresi Linear

Asumsi regresi linear meliputi hubungan linear antara prediktor dan outcome, independensi observasi, varians residual yang relatif konstan, normalitas residual, dan tidak adanya outlier berpengaruh yang mendominasi model. Linearitas dapat diperiksa melalui scatterplot atau residual plot. Homoskedastisitas dapat diperiksa dengan plot residual terhadap fitted values. Normalitas residual dapat diperiksa dengan Q–Q plot residual. Outlier berpengaruh dapat diperiksa dengan leverage, Cook's distance, atau analisis sensitivitas.

Pelanggaran asumsi tidak selalu berarti model tidak dapat digunakan, tetapi peneliti harus memahami dampaknya. Jika hubungan tidak linear, peneliti dapat mempertimbangkan transformasi, polynomial terms, spline, atau model nonlinear. Jika varians residual tidak konstan, robust standard errors atau transformasi outcome dapat dipertimbangkan. Jika outlier berpengaruh sangat kuat, peneliti harus memeriksa apakah outlier merupakan kesalahan data atau nilai valid yang perlu dianalisis secara sensitif.



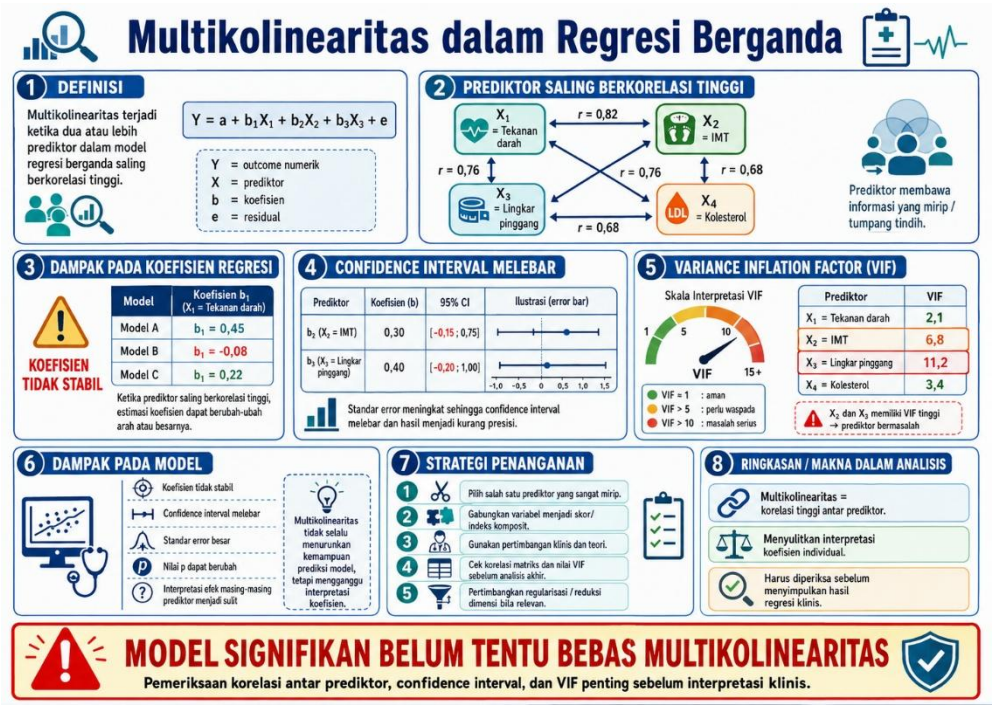
Gambar 6. 10. Pemeriksaan asumsi regresi linear

Diagram menunjukkan linearitas, residual plot, Q–Q plot residual, homoskedastisitas, outlier berpengaruh, leverage, dan Cook's distance. **Sumber:** Diadaptasi dari Ghasemi dan Zahediasl (2012), Kosin Medical Journal (2025), dan Rovetta *et al*. (2025).

6.12 Multikolinearitas dalam Regresi

Multikolinearitas terjadi ketika dua atau lebih prediktor dalam model regresi saling berkorelasi sangat kuat. Kondisi ini dapat membuat koefisien regresi menjadi tidak stabil, memperlebar confidence interval, dan menyulitkan interpretasi kontribusi masing-masing prediktor. Misalnya, berat badan dan indeks massa tubuh dapat memiliki korelasi tinggi, sehingga memasukkan keduanya dalam model yang sama perlu dipertimbangkan secara hati-hati.

Multikolinearitas dapat diperiksa melalui matriks korelasi antar prediktor, variance inflation factor, tolerance, atau perubahan koefisien ketika variabel tertentu dimasukkan atau dikeluarkan. Jika multikolinearitas tinggi, peneliti dapat memilih salah satu variabel yang paling relevan, menggabungkan variabel, menggunakan skor komposit, atau menerapkan pendekatan regularisasi pada model prediksi tertentu. Keputusan harus didasarkan pada tujuan analisis, apakah model digunakan untuk interpretasi etiologis atau prediksi.



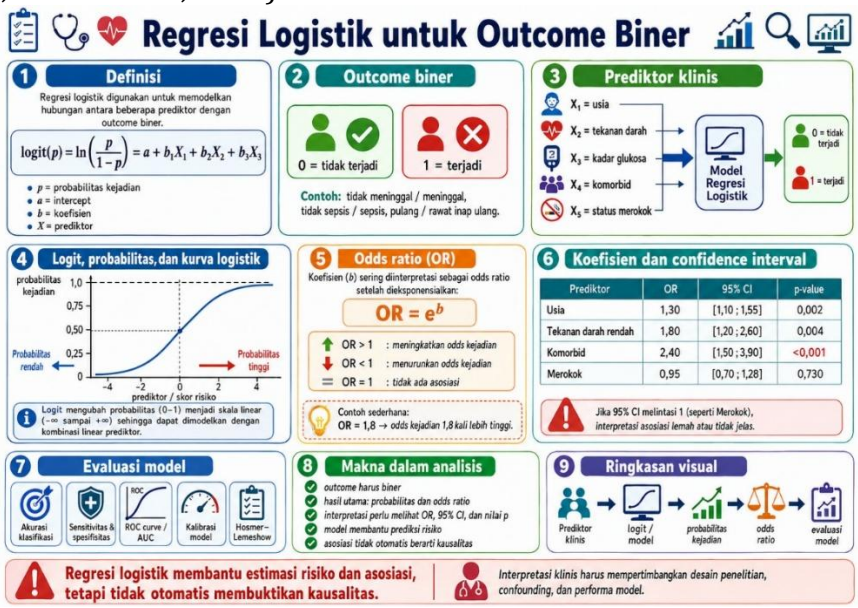
Gambar 6. 11. Multikolinearitas dalam regresi berganda

Gambar menunjukkan prediktor yang saling berkorelasi tinggi, koefisien tidak stabil, confidence interval melebar, variance inflation factor, dan strategi penanganan. **Sumber:** Diadaptasi dari Kosin Medical Journal (2025), Collins *et al*. (2024), dan prinsip model regresi klinis.

6.13 Regresi Logistik

Regresi logistik digunakan ketika outcome bersifat biner atau dikotom, seperti sakit/tidak sakit, meninggal/hidup, respons/tidak respons, atau komplikasi/tidak komplikasi. Regresi logistik memperkirakan log odds outcome berdasarkan satu atau lebih prediktor. Hasil regresi logistik biasanya dilaporkan dalam bentuk odds ratio, confidence interval, dan nilai p. Zabor *et al*. (2022) menjelaskan bahwa regresi logistik umum digunakan untuk menilai hubungan antara satu atau lebih variabel independen dengan outcome biner, tetapi kesalahan dalam pembangunan model dapat mengganggu validitas interpretasi (Zabor *et al*., 2022).

Regresi logistik tidak memerlukan normalitas outcome karena outcome biner memang tidak berdistribusi normal. Namun, model ini tetap memiliki asumsi dan kebutuhan metodologis, seperti independensi observasi, hubungan linear antara prediktor kontinu dan logit outcome, tidak adanya multikolinearitas berlebihan, jumlah kejadian yang memadai, serta evaluasi performa model. Logistic regression banyak digunakan dalam clinical risk prediction, tetapi model prediksi harus divalidasi agar tidak hanya bekerja baik pada data pengembangan (Hua, 2025; Collins *et al*., 2024).



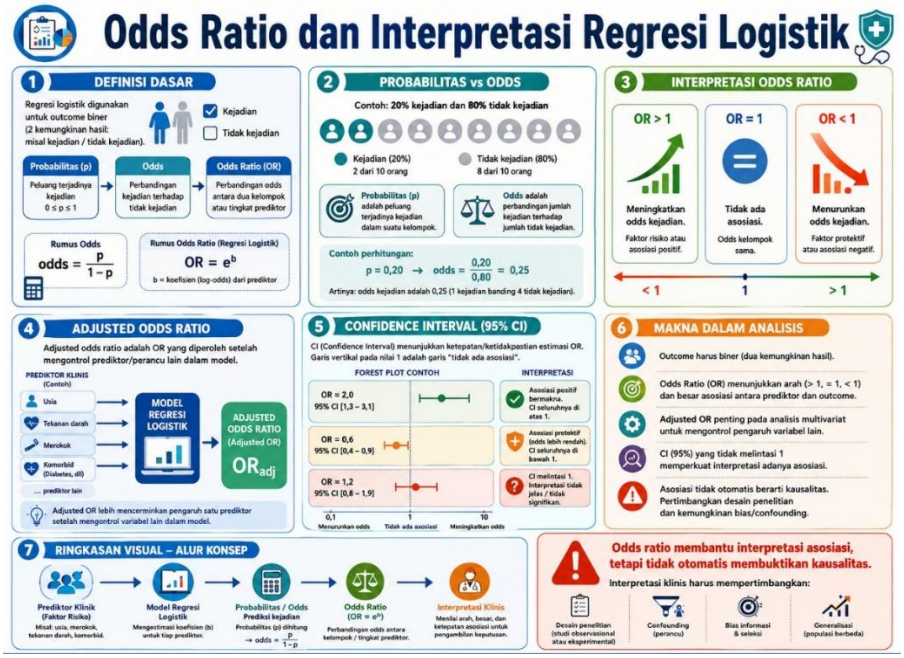
Gambar 6. 12. Regresi logistik untuk outcome biner

Diagram menunjukkan outcome biner, prediktor klinis, logit, probabilitas kejadian, odds ratio, confidence interval, dan evaluasi model. **Sumber:** Diadaptasi dari Zabor *et al*. (2022), Hua (2025), Kosin Medical Journal (2025), dan Collins *et al*. (2024).

6.14 Odds Ratio dalam Regresi Logistik

Odds ratio menunjukkan perbandingan odds kejadian outcome antara dua kelompok atau untuk setiap kenaikan satu unit prediktor. Odds berbeda dari probabilitas, karena odds adalah perbandingan peluang kejadian dengan peluang tidak kejadian. Egbuchulem *et al*. (2025) menjelaskan bahwa odds ratio merupakan ukuran asosiasi yang membandingkan odds penyakit atau kejadian pada kelompok terpapar dengan odds pada kelompok tidak terpapar (Egbuchulem *et al*., 2025). Jika odds ratio lebih dari 1, prediktor dikaitkan dengan odds outcome yang lebih tinggi; jika kurang dari 1, prediktor dikaitkan dengan odds outcome yang lebih rendah.

Interpretasi odds ratio harus dilakukan secara hati-hati, terutama jika outcome sering terjadi. Pada outcome yang jarang, odds ratio dapat mendekati risk ratio, tetapi pada outcome yang umum, odds ratio dapat tampak lebih besar dari risk ratio. Dalam regresi logistik berganda, odds ratio disebut adjusted odds ratio jika model mengontrol variabel lain. Pelaporan odds ratio sebaiknya disertai confidence interval dan nilai p agar pembaca dapat menilai besar asosiasi dan ketidakpastiannya.



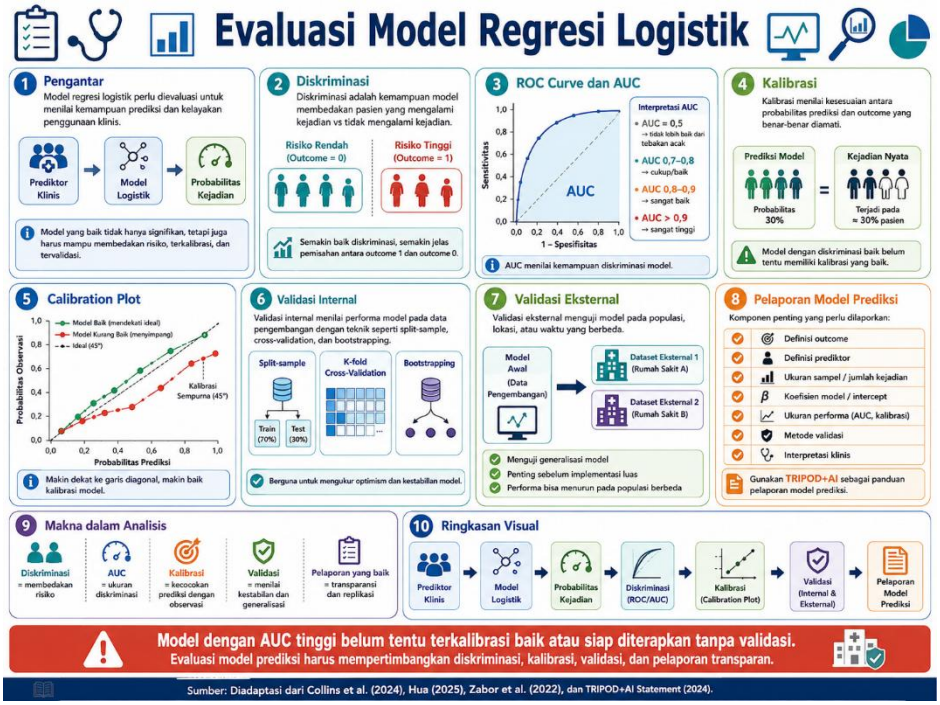
Gambar 6. 13. Odds ratio dan interpretasi regresi logistic

Gambar menunjukkan odds, probabilitas, odds ratio lebih dari 1, odds ratio kurang dari 1, adjusted odds ratio, dan confidence interval. **Sumber:** Diadaptasi dari Egbuchulem *et al*. (2025), Zabor *et al*. (2022), dan Kosin Medical Journal (2025).

6.15 Evaluasi Model Regresi Logistik

Model regresi logistik perlu dievaluasi berdasarkan kalibrasi, diskriminasi, dan kegunaan klinis. Diskriminasi menunjukkan kemampuan model membedakan subjek yang mengalami outcome dari yang tidak mengalami outcome. Area under the ROC curve sering digunakan sebagai ukuran diskriminasi. Kalibrasi menunjukkan kesesuaian antara probabilitas yang diprediksi model dan kejadian aktual. Model dapat memiliki diskriminasi baik tetapi kalibrasi buruk, sehingga keduanya perlu diperiksa.

Validasi model penting untuk mengetahui apakah model bekerja baik pada data baru. Validasi internal dapat dilakukan dengan bootstrap, cross-validation, atau split-sample, sedangkan validasi eksternal menggunakan dataset berbeda. TRIPOD+AI 2024 memperbarui pedoman pelaporan model prediksi klinis yang menggunakan regresi maupun machine learning, dengan tujuan meningkatkan transparansi laporan pengembangan dan evaluasi model (Collins *et al*., 2024).



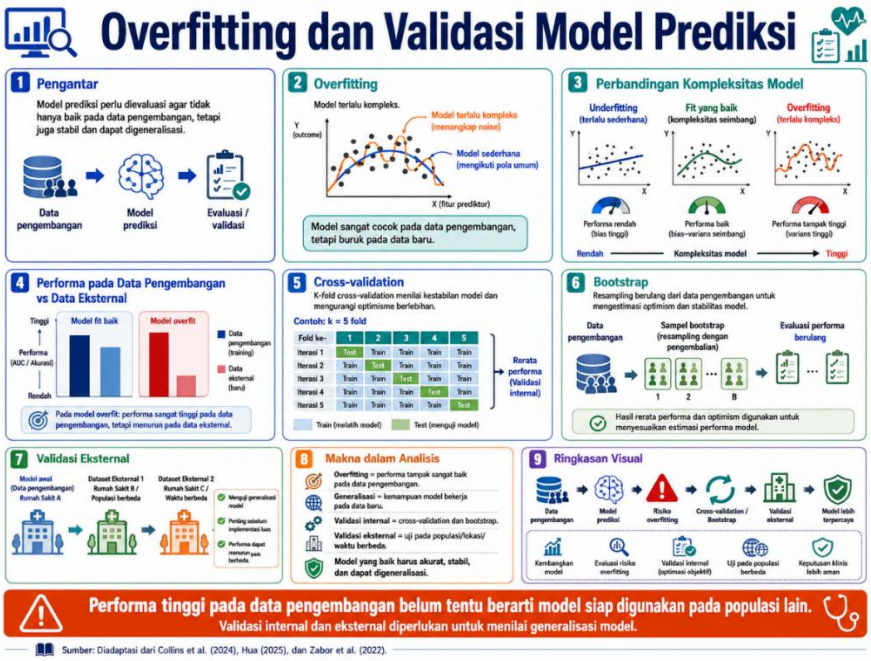
Gambar 6. 14. Evaluasi model regresi logistik

Diagram menunjukkan diskriminasi, ROC curve, AUC, kalibrasi, calibration plot, validasi internal, validasi eksternal, dan pelaporan model prediksi. **Sumber:** Diadaptasi dari Collins *et al*. (2024), Hua (2025), Zabor *et al*. (2022), dan TRIPOD+AI Statement (2024).

6.16 Overfitting dan Validasi Model

Overfitting terjadi ketika model terlalu menyesuaikan diri dengan data sampel sehingga performanya tampak baik pada data pelatihan tetapi buruk pada data baru. Overfitting sering terjadi ketika jumlah prediktor terlalu banyak dibandingkan jumlah observasi atau jumlah kejadian outcome. Dalam regresi logistik, jumlah kejadian yang terlalu sedikit untuk jumlah prediktor dapat menghasilkan koefisien yang tidak stabil. Overfitting juga dapat terjadi pada model machine learning jika validasi tidak memadai.

Pencegahan overfitting dapat dilakukan dengan perencanaan ukuran sampel yang memadai, pembatasan jumlah prediktor, pemilihan variabel berdasarkan teori klinis, regularisasi, shrinkage, cross-validation, bootstrap, dan validasi eksternal. Dalam model prediksi klinis, tujuan bukan hanya memperoleh nilai p signifikan, tetapi menghasilkan model yang akurat, terkalibrasi, dapat diterapkan, dan berguna bagi keputusan klinis. Pedoman TRIPOD+AI menekankan perlunya pelaporan transparan pada model prediksi yang menggunakan regresi atau machine learning (Collins *et al*., 2024).



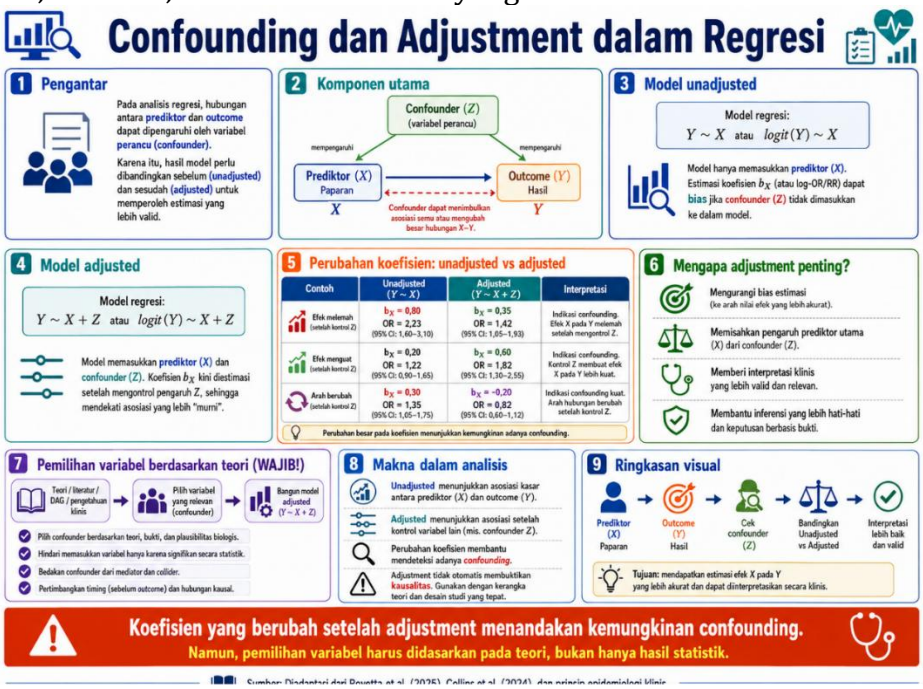
Gambar 6. 15. Overfitting dan validasi model prediksi

Gambar menunjukkan model terlalu kompleks, performa tinggi pada data pengembangan, performa menurun pada data eksternal, cross-validation, bootstrap, dan validasi eksternal. **Sumber:** Diadaptasi dari Collins *et al*. (2024), Hua (2025), dan Zabor *et al*. (2022).

6.17 Confounding dan Adjustment dalam Regresi

Confounding terjadi ketika hubungan antara prediktor dan outcome dipengaruhi oleh variabel ketiga yang berkaitan dengan keduanya. Regresi berganda dapat digunakan untuk melakukan adjustment terhadap confounder yang diukur. Misalnya, hubungan antara merokok dan penyakit periodontal dapat diadjust terhadap usia, jenis kelamin, diabetes, dan status sosial ekonomi. Setelah adjustment, koefisien prediktor utama dapat berubah, yang menunjukkan bahwa sebagian asosiasi awal dipengaruhi oleh confounder.

Namun, adjustment tidak otomatis benar jika variabel yang dimasukkan tidak dipilih berdasarkan teori atau diagram kausal yang tepat. Memasukkan mediator sebagai confounder dapat mengubah interpretasi efek. Memasukkan terlalu banyak variabel pada sampel kecil dapat menyebabkan overfitting. Oleh karena itu, pemilihan variabel dalam model harus didasarkan pada pertanyaan penelitian, pengetahuan klinis, literatur, dan struktur kausal yang masuk akal.



Gambar 6. 16. Confounding dan adjustment dalam regresi

Diagram menunjukkan prediktor, outcome, confounder, model unadjusted, model adjusted, perubahan koefisien, dan pentingnya pemilihan variabel berdasarkan teori. **Sumber:** Diadaptasi dari Rovetta *et al*. (2025), Collins *et al*. (2024), dan prinsip epidemiologi klinis.

6.18 Pelaporan Hasil Korelasi dan Regresi

Pelaporan korelasi harus mencakup jenis korelasi yang digunakan, koefisien korelasi, confidence interval, nilai p, jumlah sampel, dan alasan pemilihan Pearson, Spearman, atau Kendall. Pelaporan regresi linear harus mencakup koefisien regresi, standard error atau confidence interval, nilai p, ukuran kecocokan model seperti R-squared, pemeriksaan asumsi, dan interpretasi klinis. Pelaporan regresi logistik harus mencakup odds ratio, confidence interval, nilai p, jumlah kejadian, variabel dalam model, metode pemilihan variabel, dan evaluasi performa model bila digunakan untuk prediksi.

Untuk model prediksi, pelaporan perlu mengikuti pedoman yang lebih ketat karena model dapat digunakan untuk keputusan klinis. TRIPOD+AI memberikan panduan pelaporan untuk studi yang mengembangkan atau mengevaluasi model prediksi klinis berbasis regresi atau machine learning (Collins *et al*., 2024). Pelaporan yang transparan membantu pembaca menilai risiko bias, generalisasi, validasi, dan potensi penerapan model.



Gambar 6. 17. Pelaporan korelasi dan regresi

Gambar menunjukkan koefisien korelasi, koefisien regresi, odds ratio, confidence interval, nilai p, R-squared, AUC, kalibrasi, dan pelaporan transparan model. **Sumber:** Diadaptasi dari Collins *et al*. (2024), Wisniewski (2024), Bujang (2024), Zabor *et al*. (2022), dan Rovetta *et al*. (2025).

6.19 Kesalahan Umum dalam Analisis Korelasi dan Regresi

Kesalahan umum pertama adalah menyimpulkan kausalitas hanya berdasarkan korelasi. Kesalahan umum kedua adalah menggunakan Pearson tanpa memeriksa scatterplot, linearitas, dan outlier. Kesalahan umum ketiga adalah menggunakan regresi linear untuk outcome biner, padahal regresi logistik lebih sesuai untuk outcome dikotom. Kesalahan umum keempat adalah menafsirkan odds ratio sebagai risk ratio tanpa memperhatikan frekuensi outcome.

Kesalahan lain adalah memasukkan terlalu banyak prediktor pada ukuran sampel kecil, mengabaikan multikolinearitas, tidak memeriksa residual, tidak melaporkan confidence interval, dan hanya menekankan nilai p. Dalam model prediksi, kesalahan serius mencakup tidak melakukan validasi, melaporkan performa hanya pada data pengembangan, dan tidak mengikuti pedoman pelaporan. TRIPOD+AI menekankan transparansi pelaporan model prediksi klinis agar pembaca dapat menilai kualitas dan kegunaan model (Collins *et al*., 2024).

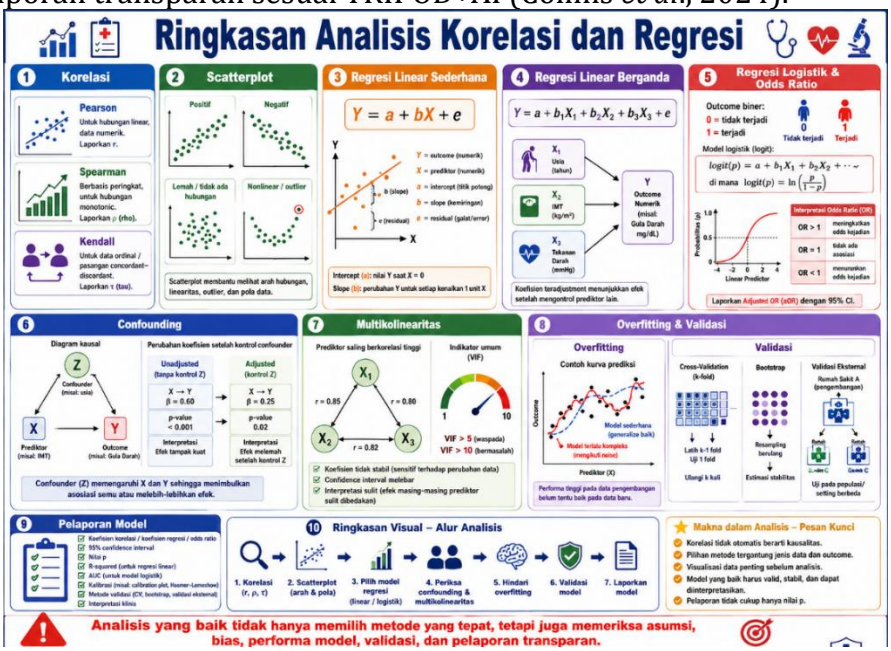


Gambar 6. 18. Kesalahan umum korelasi dan regresi

Diagram menunjukkan korelasi dianggap kausal, Pearson tanpa scatterplot, regresi linear untuk outcome biner, overfitting, multikolinearitas, dan pelaporan tanpa validasi. **Sumber:** Diadaptasi dari Ahmed (2025), Zabor *et al*. (2022), Collins *et al*. (2024), dan Rovetta *et al*. (2025).

Bab ini menjelaskan bahwa korelasi dan regresi merupakan metode utama untuk menilai hubungan antarvariabel dalam penelitian kuantitatif. Korelasi Pearson digunakan untuk hubungan linear dua variabel numerik, sedangkan Spearman dan Kendall lebih sesuai untuk data ordinal, hubungan monotonic, atau kondisi data yang tidak memenuhi asumsi Pearson. Korelasi menunjukkan arah dan kekuatan asosiasi, tetapi tidak membuktikan kausalitas. Scatterplot harus diperiksa sebelum interpretasi korelasi karena outlier dan pola nonlinear dapat mengubah makna koefisien.

Regresi linear digunakan untuk outcome numerik kontinu, sedangkan regresi logistik digunakan untuk outcome biner. Regresi berganda dapat mengontrol confounding, tetapi tetap memerlukan pemilihan variabel yang tepat, pemeriksaan asumsi, dan ukuran sampel memadai. Odds ratio adalah ukuran asosiasi utama pada regresi logistik, tetapi harus dibedakan dari probabilitas dan risk ratio. Model prediksi klinis memerlukan evaluasi diskriminasi, kalibrasi, validasi, dan pelaporan transparan sesuai TRIPOD+AI (Collins *et al*., 2024).



Gambar 6. 19. Ringkasan analisis korelasi dan regresi

Diagram merangkum Pearson, Spearman, Kendall, scatterplot, regresi linear, regresi linear berganda, regresi logistik, odds ratio, confounding, multikolinearitas, overfitting, validasi, dan pelaporan model. **Sumber:** Diadaptasi dari Wisniewski (2024), Bujang (2024), Zabor *et al*. (2022), Egbuchulem *et al*. (2025), dan Collins *et al*. (2024).

Review Questions

1. **Apa perbedaan korelasi dan regresi?**
   Jawaban: Korelasi menilai arah dan kekuatan hubungan antarvariabel, sedangkan regresi memodelkan atau memprediksi outcome berdasarkan satu atau lebih prediktor.
2. **Kapan korelasi Pearson digunakan?**
   Jawaban: Pearson digunakan ketika dua variabel numerik memiliki hubungan linear dan tidak didominasi oleh outlier ekstrem.
3. **Kapan korelasi Spearman lebih sesuai daripada Pearson?**
   Jawaban: Spearman lebih sesuai untuk data ordinal, data tidak normal, hubungan monotonic, atau data dengan outlier yang mengganggu Pearson.
4. **Apa arti koefisien korelasi positif?**
   Jawaban: Koefisien positif berarti nilai satu variabel cenderung meningkat ketika variabel lain meningkat.
5. **Mengapa korelasi tidak membuktikan kausalitas?**
   Jawaban: Korelasi dapat dipengaruhi oleh confounding, reverse causation, bias, atau kebetulan statistik.
6. **Kapan regresi linear digunakan?**
   Jawaban: Regresi linear digunakan ketika outcome bersifat numerik kontinu dan hubungan dengan prediktor dapat dimodelkan secara linear.
7. **Apa fungsi regresi linear berganda?**
   Jawaban: Regresi linear berganda digunakan untuk menjelaskan outcome numerik dengan beberapa prediktor dan dapat mengontrol confounding yang diukur.
8. **Kapan regresi logistik digunakan?**
   Jawaban: Regresi logistik digunakan ketika outcome bersifat biner atau dikotom, seperti sakit/tidak sakit atau komplikasi/tidak komplikasi.
9. **Apa yang dimaksud odds ratio?**
   Jawaban: Odds ratio adalah perbandingan odds kejadian outcome antara kelompok atau untuk setiap kenaikan prediktor dalam regresi logistik.
10. **Mengapa validasi model penting?**
    Jawaban: Validasi penting untuk memastikan model tidak hanya bekerja baik pada data pengembangan, tetapi juga memiliki performa yang dapat diterapkan pada data atau populasi baru.

DAFTAR PUSTAKA

Ahmed, H. S. (2025). Correlation and causation for cardiothoracic surgeons. *Journal of Thoracic Disease / PMC-indexed article*.

Bujang, M. A. (2024). An elaboration on sample size determination for correlation analysis: A brief guide for researchers. *Restorative Dentistry & Endodontics, 49*, e21. https://doi.org/10.5395/rde.2024.49.e21

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*, Article 40. https://doi.org/10.1186/s13040-025-00465-6

Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., *et al*. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ, 385*, e078378.

Egbuchulem, K. I., *et al*. (2025). The odds ratio: A measure of strength in clinical research. *PMC-indexed article*.

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

Hua, Y. (2025). Clinical risk prediction with logistic regression: Best practices, validation techniques, and applications in medical research. *Academic Medicine & Surgery*.

Kosin Medical Journal. (2025). Common statistical methods used in medical research. *Kosin Medical Journal*.

Rovetta, A., Piretta, L., & Mansournia, M. A. (2025). p-values and confidence intervals as compatibility measures: Guidelines for interpreting statistical studies in clinical research. *The Lancet Regional Health – Southeast Asia*.

Wisniewski, S. J. (2024). Correlation coefficient, partial correlation, Spearman rank correlation, and regression analysis. In *StatPearls*. StatPearls Publishing.

Zabor, E. C., *et al*. (2022). Logistic regression in clinical studies. *International Journal of Radiation Oncology, Biology, Physics*.

BAB 7
ANALISIS DATA KATEGORIK DAN DIAGNOSTIK: TABEL KONTINGENSI, CHI-SQUARE, FISHER'S EXACT, MCNEMAR, SENSITIVITAS, SPESIFISITAS, DAN ROC CURVE

7.1 Pendahuluan

Analisis data kategorik digunakan ketika variabel penelitian berbentuk kategori, seperti jenis kelamin, status penyakit, status paparan, kelompok terapi, respons klinis, atau kejadian komplikasi (Kosin Medical Journal, 2025). Data kategorik dapat bersifat nominal ketika kategori tidak memiliki urutan, atau ordinal ketika kategori memiliki tingkatan tertentu (Kaliyadan & Kulkarni, 2019). Dalam penelitian kesehatan, analisis kategorik sering digunakan untuk membandingkan proporsi kejadian antar kelompok, menilai hubungan antara paparan dan outcome, atau mengevaluasi perubahan status sebelum dan sesudah intervensi (Kosin Medical Journal, 2025). Uji yang sering digunakan untuk data kategorik meliputi chi-square test, Fisher's exact test, McNemar test, serta ukuran asosiasi seperti risk ratio, odds ratio, dan risk difference (Kosin Medical Journal, 2025; Sundjaja *et al*., 2023).

Analisis diagnostik merupakan perluasan penting dari data kategorik karena banyak penelitian klinis menilai apakah suatu pemeriksaan mampu membedakan pasien sakit dan tidak sakit (Çorbacıoğlu & Aksel, 2023). Ukuran utama dalam analisis diagnostik meliputi sensitivitas, spesifisitas, nilai prediksi positif, nilai prediksi negatif, likelihood ratio, accuracy, dan area under the ROC curve (Ayad *et al*., 2026). ROC curve digunakan untuk menilai performa diskriminatif suatu tes diagnostik atau biomarker pada berbagai titik cut-off (Çorbacıoğlu & Aksel, 2023). Namun, analisis ROC harus dilakukan hati-hati karena kesalahan dalam membandingkan AUC atau memilih cut-off

dapat menghasilkan interpretasi diagnostik yang tidak tepat (He *et al*., 2025).



Gambar 7. 1. Peta konsep analisis data kategorik dan diagnostic Ilustrasi menunjukkan variabel kategorik, tabel kontingensi, chi-square test, Fisher's exact test, McNemar test, risk ratio, odds ratio, sensitivitas, spesifisitas, likelihood ratio, ROC curve, dan AUC. **Sumber:** Diadaptasi dari Kosin Medical Journal (2025), Sundjaja *et al*. (2023), Çorbacıoğlu dan Aksel (2023), Ayad *et al*. (2026), serta He *et al*. (2025).

7.2 Jenis Data Kategorik

Data kategorik adalah data yang diklasifikasikan ke dalam kelompok atau kategori tertentu (Kaliyadan & Kulkarni, 2019). Data kategorik nominal tidak memiliki urutan alami, seperti jenis kelamin, golongan darah, status merokok, jenis terapi, atau diagnosis penyakit (Kaliyadan & Kulkarni, 2019). Data kategorik ordinal memiliki urutan, seperti derajat nyeri ringan-sedang-berat, stadium penyakit, tingkat kepuasan, atau skala respons terapi (Kaliyadan & Kulkarni, 2019). Pemilihan analisis untuk data kategorik harus mempertimbangkan apakah kategori bersifat nominal, ordinal, independen, atau berpasangan (Kosin Medical Journal, 2025).

Data kategorik biasanya disajikan dengan frekuensi dan persentase (Kaliyadan & Kulkarni, 2019). Jika peneliti membandingkan dua variabel kategorik independen, tabel kontingensi dapat digunakan untuk merangkum jumlah observasi pada setiap kombinasi kategori (Kosin Medical Journal, 2025). Jika data kategorik berasal dari subjek yang sama pada dua waktu berbeda, analisis harus memperhitungkan pasangan data, sehingga McNemar test lebih tepat daripada chi-square biasa pada tabel 2×2 berpasangan (Sundjaja *et al*., 2023). Kesalahan membedakan data independen dan berpasangan dapat menyebabkan pemilihan uji yang tidak sesuai (Sundjaja *et al*., 2023).



Gambar 7. 2. Klasifikasi data kategorik

Diagram menunjukkan data nominal, ordinal, dikotom, politom, independen, berpasangan, serta hubungan masing-masing dengan pilihan uji statistik. **Sumber:** Diadaptasi dari Kaliyadan dan Kulkarni (2019), Kosin Medical Journal (2025), serta Sundjaja *et al*. (2023).

7.3 Tabel Kontingensi

Tabel kontingensi adalah tabel yang menyajikan frekuensi observasi berdasarkan kombinasi dua atau lebih variabel kategorik (Kosin Medical Journal, 2025). Tabel 2×2 sering digunakan ketika masing-masing variabel memiliki dua kategori, seperti paparan ya/tidak dan

penyakit ya/tidak (Kosin Medical Journal, 2025). Tabel r×c digunakan ketika salah satu atau kedua variabel memiliki lebih dari dua kategori (Kosin Medical Journal, 2025). Tabel kontingensi menjadi dasar untuk menghitung proporsi, risk ratio, odds ratio, chi-square test, Fisher's exact test, dan McNemar test sesuai desain data (Kosin Medical Journal, 2025).

Tabel kontingensi harus menampilkan jumlah observasi aktual, bukan hanya persentase (Kosin Medical Journal, 2025). Persentase dapat dihitung berdasarkan baris, kolom, atau total, sehingga peneliti harus menjelaskan dasar persentase yang digunakan (Kosin Medical Journal, 2025). Pada studi kohort atau uji klinis, persentase berdasarkan kelompok paparan atau kelompok terapi sering digunakan untuk membandingkan risiko outcome (Kosin Medical Journal, 2025). Pada studi kasus-kontrol, odds ratio lebih sering digunakan karena sampling dilakukan berdasarkan status outcome (Kosin Medical Journal, 2025).

Tabel 7. 1. Contoh tabel kontingensi 2×2

Kelompok	Outcome positif	Outcome negatif	Total
Terpapar / intervensi	a	b	a + b
Tidak terpapar / kontrol	c	d	c + d
Total	a + c	b + d	n

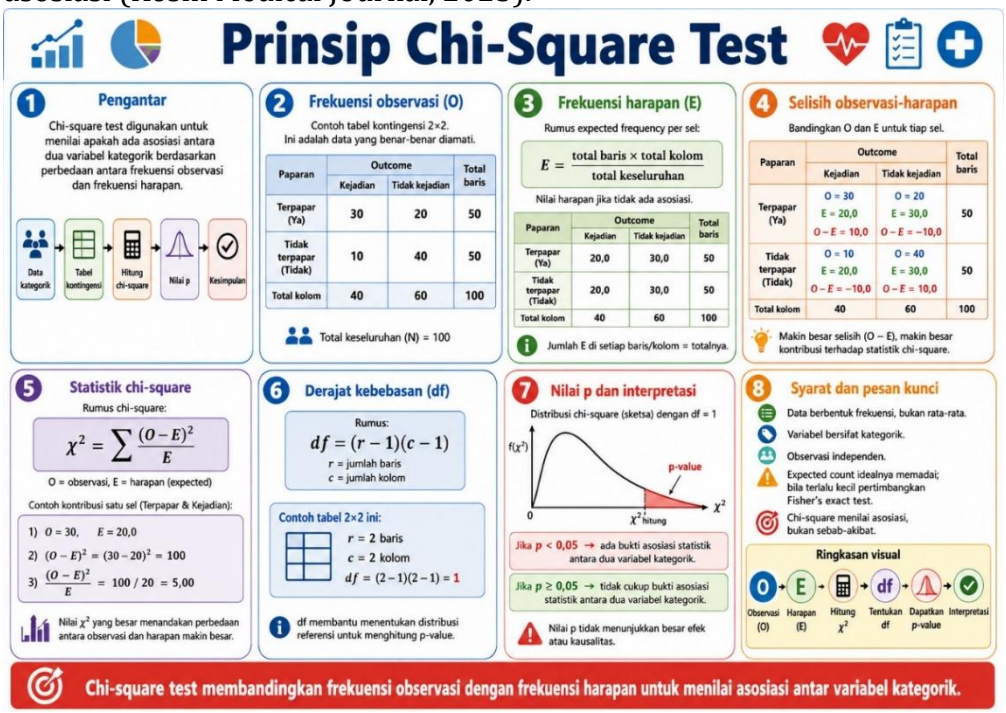
Sumber tabel: Diadaptasi dari Kosin Medical Journal (2025) dan prinsip epidemiologi klinis.

7.4 Chi-Square Test

Chi-square test digunakan untuk menilai apakah terdapat hubungan atau perbedaan proporsi antara dua variabel kategorik independen (Kosin Medical Journal, 2025). Uji ini membandingkan frekuensi observasi dengan frekuensi harapan yang diasumsikan terjadi bila tidak ada hubungan antara variabel (Kosin Medical Journal, 2025). Jika perbedaan antara frekuensi observasi dan frekuensi harapan cukup besar, nilai chi-square meningkat dan nilai p dapat menjadi lebih kecil (Kosin Medical Journal, 2025). Chi-square test banyak digunakan dalam penelitian kesehatan untuk membandingkan proporsi kejadian, respons terapi, status penyakit, atau komplikasi antar kelompok (Kosin Medical Journal, 2025). (

Chi-square test memiliki asumsi penting, terutama independensi observasi dan kecukupan expected count dalam sel tabel (Chicco *et al*., 2025). Jika expected count terlalu rendah, Fisher's exact test lebih sesuai digunakan, terutama pada tabel 2×2 dengan sampel kecil (Chicco *et al*., 2025). Chi-square test tidak menunjukkan besar asosiasi, sehingga

hasilnya sebaiknya dilengkapi dengan ukuran efek seperti risk ratio, odds ratio, risk difference, atau Cramer's V sesuai konteks (Kosin Medical Journal, 2025). Pelaporan chi-square sebaiknya mencakup frekuensi, persentase, statistik chi-square, derajat kebebasan, nilai p, dan ukuran asosiasi (Kosin Medical Journal, 2025).



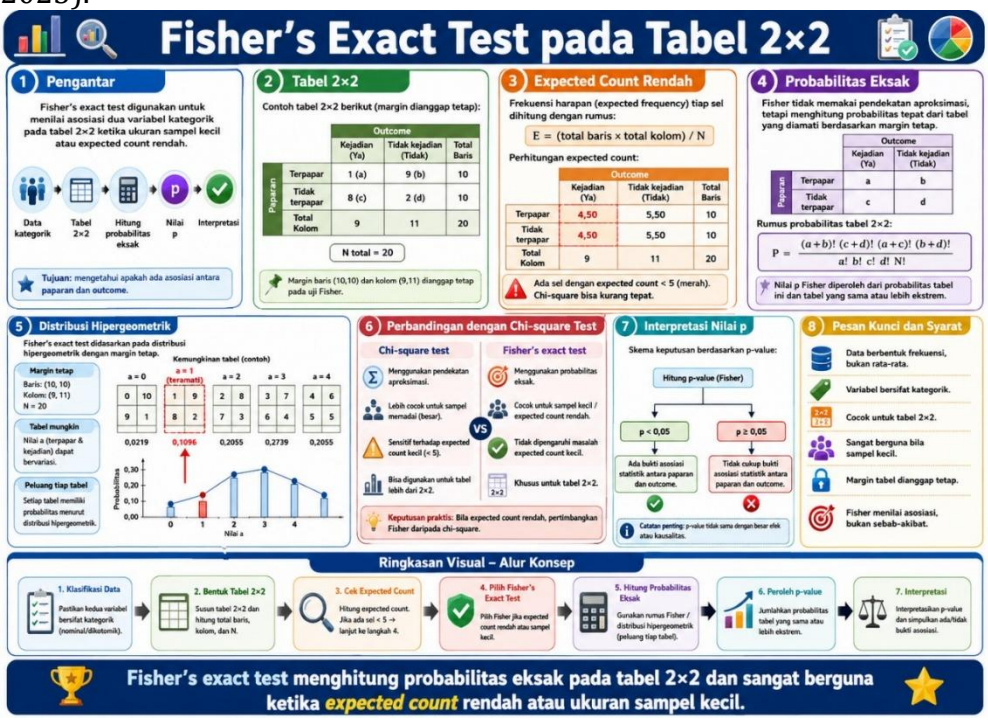
Gambar 7. 3. Prinsip chi-square test

Gambar menunjukkan frekuensi observasi, frekuensi harapan, selisih observasi-harapan, statistik chi-square, derajat kebebasan, dan nilai p. **Sumber:** Diadaptasi dari Kosin Medical Journal (2025) dan Chicco *et al*. (2025).

7.5 Fisher's Exact Test

Fisher's exact test digunakan untuk menganalisis tabel kontingensi ketika ukuran sampel kecil atau expected count rendah (Chicco *et al*., 2025). Uji ini menghitung probabilitas konfigurasi tabel secara eksak berdasarkan distribusi hipergeometrik, sehingga tidak bergantung pada pendekatan asimtotik seperti chi-square (Chicco *et al*., 2025). Fisher's exact test sangat sering digunakan pada tabel 2×2 ketika data jarang atau jumlah kejadian sedikit (Chicco *et al*., 2025). Dalam penelitian klinis, uji ini berguna ketika kejadian komplikasi, efek samping, atau outcome langka memiliki jumlah sel yang kecil (Kosin Medical Journal, 2025).

Fisher's exact test tidak boleh dianggap hanya sebagai "uji cadangan" tanpa pertimbangan desain penelitian (Chicco *et al*., 2025). Jika sampel besar dan expected count memadai, chi-square test biasanya cukup digunakan (Kosin Medical Journal, 2025). Jika data bersifat berpasangan, Fisher's exact test biasa tidak menggantikan McNemar test karena struktur dependensi data berbeda (Sundjaja *et al*., 2023). Oleh karena itu, pemilihan Fisher's exact test harus mempertimbangkan ukuran sampel, expected count, dan independensi observasi (Chicco *et al*., 2025).



Gambar 7. 4. Fisher's exact test pada tabel 2×2

Diagram menunjukkan tabel 2×2 dengan expected count rendah, probabilitas eksak, distribusi hipergeometrik, dan perbandingan dengan chi-square test.

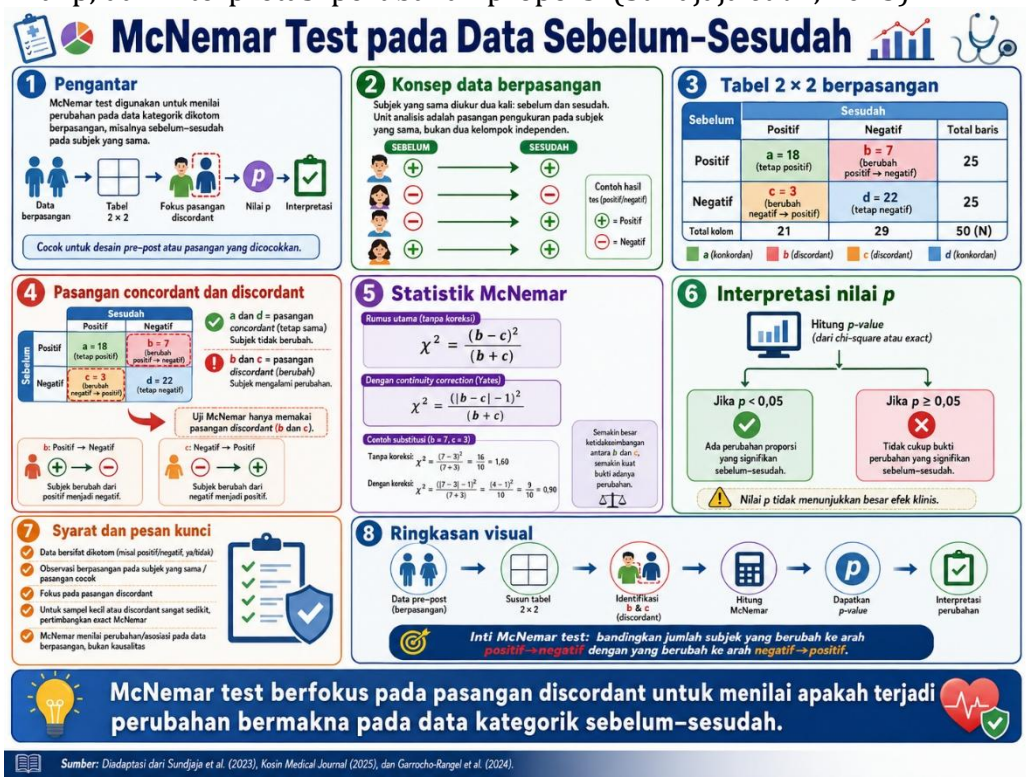
Sumber: Diadaptasi dari Chicco *et al*. (2025) dan Kosin Medical Journal (2025).

7.6 McNemar Test untuk Data Berpasangan

McNemar test digunakan untuk menganalisis data nominal dikotom berpasangan pada tabel 2×2 (Sundjaja *et al*., 2023). Uji ini sering digunakan pada desain sebelum-sesudah, matched-pair study, atau perbandingan dua metode diagnostik pada subjek yang sama (Sundjaja *et al*., 2023). McNemar test menilai marginal homogeneity, yaitu apakah

proporsi kategori positif berbeda antara dua kondisi berpasangan (Sundjaja *et al*., 2023). Dalam contoh klinis, McNemar test dapat digunakan untuk menilai apakah status positif-negatif pasien berubah setelah intervensi (Kosin Medical Journal, 2025).

McNemar test berfokus pada pasangan discordant, yaitu subjek yang berubah dari positif ke negatif atau negatif ke positif (Sundjaja *et al*., 2023). Subjek yang tetap positif atau tetap negatif tidak menjadi sumber utama informasi untuk menguji perubahan proporsi (Sundjaja *et al*., 2023). Jika jumlah pasangan discordant kecil, versi exact McNemar dapat dipertimbangkan (Garrocho-Rangel *et al*., 2024). Pelaporan McNemar test sebaiknya menyajikan tabel 2×2 berpasangan, jumlah discordant pairs, nilai p, dan interpretasi perubahan proporsi (Sundjaja *et al*., 2023).



Gambar 7. 5. McNemar test pada data sebelum–sesudah

Gambar menunjukkan tabel 2×2 berpasangan, subjek tetap positif, tetap negatif, berubah positif ke negatif, berubah negatif ke positif, dan pasangan discordant sebagai dasar uji. **Sumber:** Diadaptasi dari Sundjaja *et al*. (2023), Kosin Medical Journal (2025), dan Garrocho-Rangel *et al*. (2024).

7.7 Risk Difference, Risk Ratio, dan Odds Ratio

Risk difference menunjukkan selisih risiko outcome antara kelompok terpapar dan tidak terpapar (Kosin Medical Journal, 2025). Risk ratio menunjukkan perbandingan risiko outcome antara dua kelompok, sehingga nilai lebih dari 1 menunjukkan risiko lebih tinggi pada kelompok terpapar (Kosin Medical Journal, 2025). Odds ratio menunjukkan perbandingan odds outcome antara dua kelompok atau antara tingkat prediktor dalam model logistik (Egbuchulem *et al*., 2025). Odds ratio sering digunakan pada studi kasus-kontrol dan regresi logistik, sedangkan risk ratio lebih intuitif pada studi kohort atau uji klinis (Kosin Medical Journal, 2025).

Interpretasi odds ratio harus hati-hati karena odds tidak sama dengan probabilitas (Egbuchulem *et al*., 2025). Jika outcome jarang, odds ratio dapat mendekati risk ratio, tetapi jika outcome sering, odds ratio dapat tampak lebih ekstrem daripada risk ratio (Egbuchulem *et al*., 2025). Ukuran asosiasi sebaiknya dilaporkan bersama confidence interval agar pembaca dapat menilai presisi estimasi (Rovetta *et al*., 2025). Nilai p saja tidak cukup karena nilai p tidak menunjukkan besar asosiasi atau relevansi klinis (Rovetta *et al*., 2025).

Tabel 7. 2. Ukuran asosiasi pada data kategorik

Ukuran	Makna	Kapan sering digunakan
Risk difference	Selisih risiko absolut antar kelompok	Uji klinis, kohort, evaluasi dampak absolut
Risk ratio	Perbandingan risiko antar kelompok	Kohort, uji klinis
Odds ratio	Perbandingan odds antar kelompok	Kasus-kontrol, regresi logistik
Cramer's V	Kekuatan asosiasi kategori	Tabel kontingensi lebih besar

Sumber tabel: Diadaptasi dari Kosin Medical Journal (2025), Egbuchulem *et al*. (2025), dan Rovetta *et al*. (2025).

7.8 Analisis Diagnostik dan Tabel 2×2

Analisis diagnostik biasanya membandingkan hasil tes indeks dengan standar referensi untuk menentukan apakah seseorang benar-benar memiliki kondisi penyakit (Çorbacıoğlu & Aksel, 2023). Hasil tes diagnostik diklasifikasikan menjadi true positive, false positive, true negative, dan false negative (Ayad *et al*., 2026). True positive berarti tes positif pada pasien yang benar-benar sakit, sedangkan false positive

berarti tes positif pada pasien yang sebenarnya tidak sakit (Ayad *et al*., 2026). True negative berarti tes negatif pada pasien yang tidak sakit, sedangkan false negative berarti tes negatif pada pasien yang sebenarnya sakit (Ayad *et al*., 2026).

Tabel 2×2 diagnostik menjadi dasar perhitungan sensitivitas, spesifisitas, predictive values, likelihood ratios, dan accuracy (Ayad *et al*., 2026). Sensitivitas dan spesifisitas menggambarkan performa tes terhadap status penyakit, sedangkan nilai prediksi positif dan negatif dipengaruhi oleh prevalensi penyakit pada populasi yang diuji (Ayad *et al*., 2026). Oleh karena itu, nilai prediksi dari suatu tes dapat berubah ketika tes digunakan pada populasi dengan prevalensi berbeda (Ayad *et al*., 2026). Pelaporan studi diagnostik harus menjelaskan standar referensi, populasi uji, cut-off, serta ukuran performa diagnostik yang digunakan (Çorbacıoğlu & Aksel, 2023).

Tabel 7.3. Tabel 2×2 diagnostik

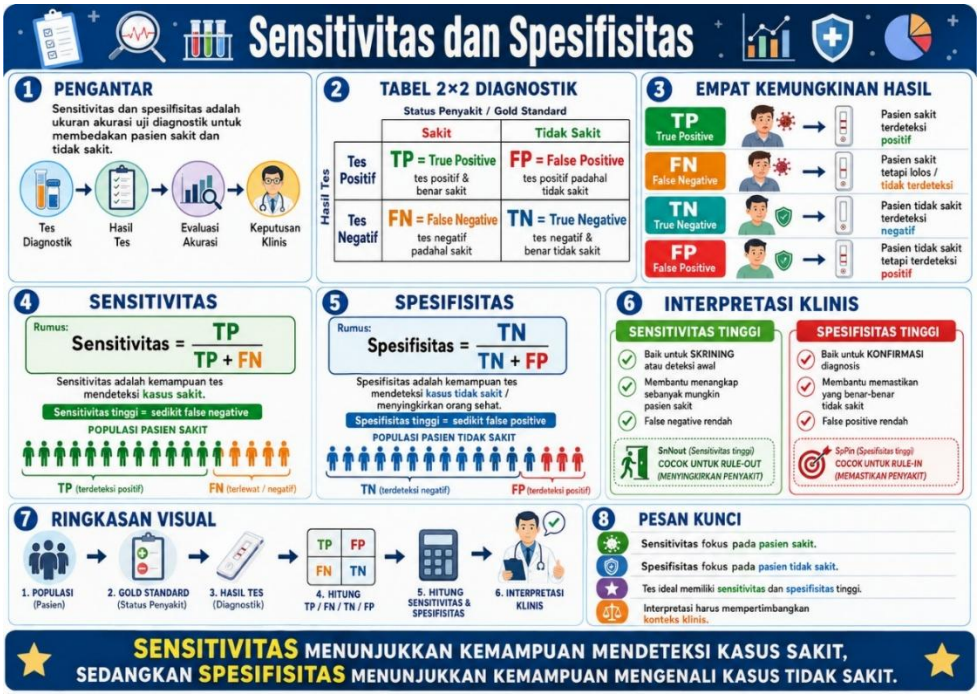
Hasil tes	Penyakit positif	Penyakit negatif
Tes positif	True positive	False positive
Tes negatif	False negative	True negative

Sumber tabel: Diadaptasi dari Ayad *et al*. (2026) dan Çorbacıoğlu dan Aksel (2023).

7.9 Sensitivitas dan Spesifisitas

Sensitivitas adalah kemampuan tes untuk mengidentifikasi pasien yang benar-benar memiliki penyakit (Ayad *et al*., 2026). Sensitivitas dihitung sebagai true positive dibagi seluruh pasien yang benar-benar sakit (Ayad *et al*., 2026). Tes dengan sensitivitas tinggi memiliki false negative rendah, sehingga berguna untuk skrining ketika kegagalan mendeteksi penyakit berisiko besar (Ayad *et al*., 2026). Namun, sensitivitas tinggi tidak otomatis berarti tes memiliki sedikit false positive (Çorbacıoğlu & Aksel, 2023).

Spesifisitas adalah kemampuan tes untuk mengidentifikasi individu yang benar-benar tidak memiliki penyakit (Ayad *et al*., 2026). Spesifisitas dihitung sebagai true negative dibagi seluruh individu yang benar-benar tidak sakit (Ayad *et al*., 2026). Tes dengan spesifisitas tinggi memiliki false positive rendah, sehingga berguna untuk konfirmasi diagnosis ketika konsekuensi diagnosis salah positif besar (Çorbacıoğlu & Aksel, 2023). Sensitivitas dan spesifisitas sering berubah berlawanan arah ketika cut-off tes digeser (Çorbacıoğlu & Aksel, 2023).



Gambar 7. 6. Sensitivitas dan spesifisitas

Gambar menunjukkan true positive, false negative, true negative, false positive, sensitivitas sebagai deteksi kasus sakit, dan spesifisitas sebagai deteksi kasus tidak sakit. **Sumber:** Diadaptasi dari Ayad *et al*. (2026) dan Çorbacıoğlu dan Aksel (2023).

7.10 Nilai Prediksi Positif dan Nilai Prediksi Negatif

Nilai prediksi positif adalah probabilitas bahwa seseorang benar-benar sakit ketika hasil tesnya positif (Ayad *et al*., 2026). Nilai prediksi negatif adalah probabilitas bahwa seseorang benar-benar tidak sakit ketika hasil tesnya negatif (Ayad *et al*., 2026). Berbeda dari sensitivitas dan spesifisitas, nilai prediksi sangat dipengaruhi oleh prevalensi penyakit dalam populasi yang diuji (Ayad *et al*., 2026). Jika prevalensi penyakit meningkat, nilai prediksi positif biasanya meningkat, sedangkan nilai prediksi negatif dapat menurun (Ayad *et al*., 2026).

Dalam praktik klinis, nilai prediksi sering lebih intuitif bagi dokter dan pasien karena menjawab pertanyaan setelah hasil tes diketahui (Çorbacıoğlu & Aksel, 2023). Namun, nilai prediksi dari satu studi tidak boleh langsung diterapkan pada populasi lain dengan prevalensi penyakit yang sangat berbeda (Ayad *et al*., 2026). Oleh karena itu, interpretasi nilai prediksi harus selalu menyebutkan konteks populasi dan prevalensi dasar (Ayad *et al*., 2026). Pelaporan diagnostik yang baik sebaiknya menyajikan

sensitivitas, spesifisitas, nilai prediksi, likelihood ratio, serta confidence interval (Çorbacıoğlu & Aksel, 2023).



Gambar 7. 7. Nilai prediksi dan pengaruh prevalensi

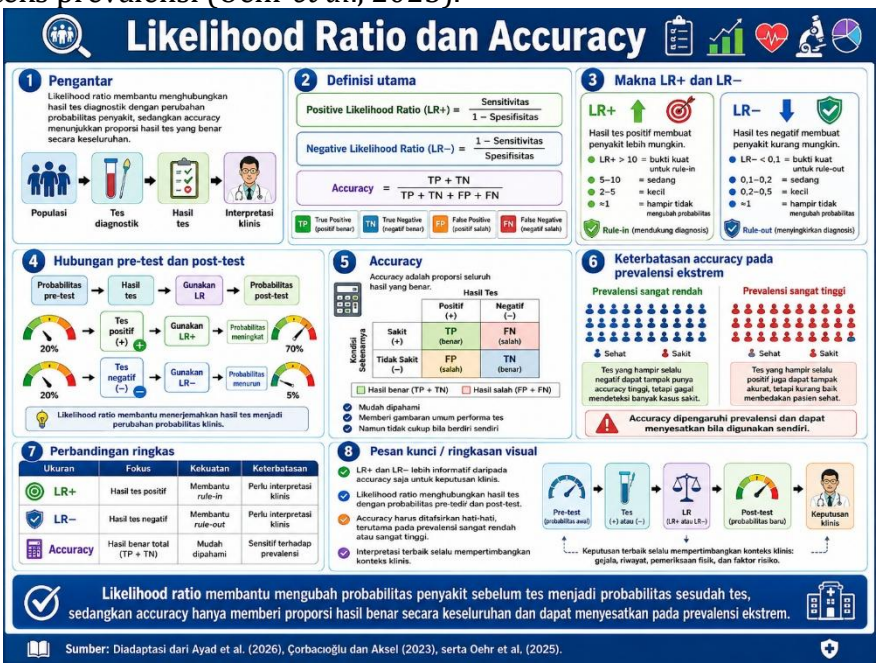
Diagram menunjukkan hubungan prevalensi penyakit dengan nilai prediksi positif dan nilai prediksi negatif pada tes diagnostik. **Sumber:** Diadaptasi dari Ayad *et al*. (2026) dan Çorbacıoğlu dan Aksel (2023).

7.11 Likelihood Ratio dan Accuracy

Positive likelihood ratio menunjukkan seberapa besar kemungkinan hasil tes positif ditemukan pada pasien sakit dibandingkan individu tidak sakit (Ayad *et al*., 2026). Negative likelihood ratio menunjukkan seberapa besar kemungkinan hasil tes negatif ditemukan pada pasien sakit dibandingkan individu tidak sakit (Ayad *et al*., 2026). Likelihood ratio berguna karena dapat menghubungkan probabilitas pre-test dengan probabilitas post-test dalam penalaran diagnostik (Çorbacıoğlu & Aksel, 2023). Positive likelihood ratio yang tinggi memperkuat kemungkinan diagnosis, sedangkan negative likelihood ratio yang rendah membantu menyingkirkan diagnosis (Çorbacıoğlu & Aksel, 2023).

Accuracy adalah proporsi seluruh hasil tes yang benar, yaitu true positive dan true negative dibandingkan total subjek (Ayad *et al*., 2026). Accuracy mudah dipahami, tetapi dapat menyesatkan bila prevalensi

penyakit sangat rendah atau sangat tinggi (Oehr *et al*., 2025). Misalnya, pada penyakit langka, tes yang selalu menyatakan "negatif" dapat memiliki accuracy tinggi tetapi tidak berguna untuk mendeteksi kasus positif (Oehr *et al*., 2025). Oleh karena itu, accuracy sebaiknya tidak digunakan sendirian tanpa sensitivitas, spesifisitas, predictive values, dan konteks prevalensi (Oehr *et al*., 2025).



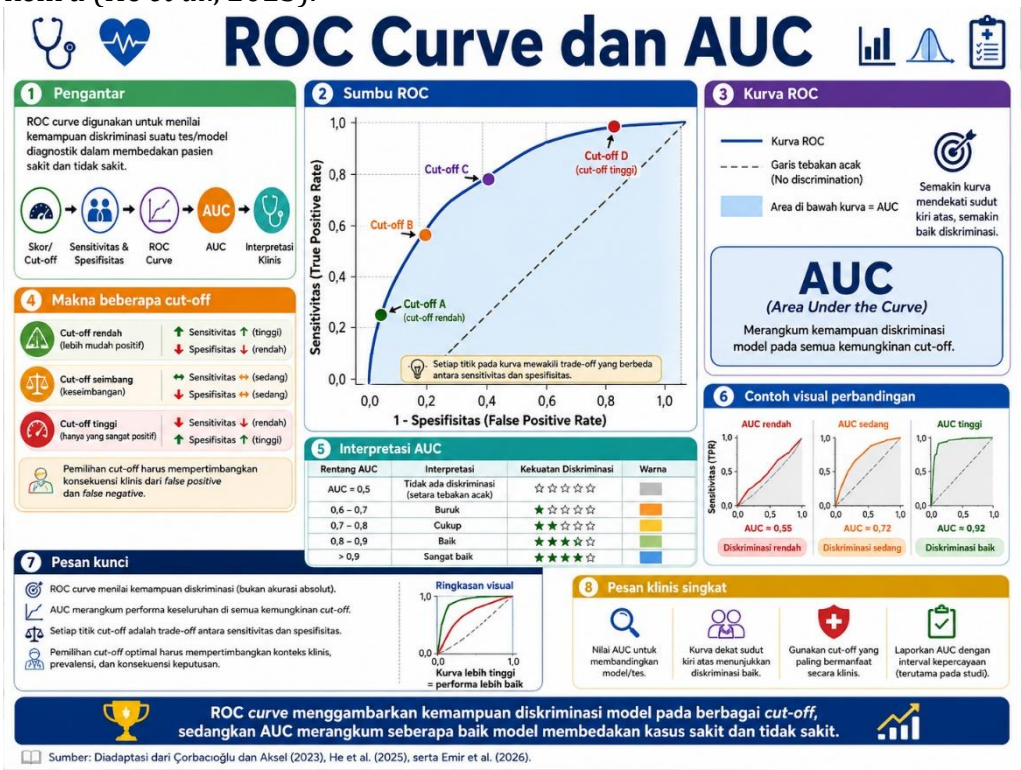
Gambar 7. 8. Likelihood ratio dan accuracy

Gambar menunjukkan positive likelihood ratio, negative likelihood ratio, accuracy, keterbatasan accuracy pada prevalensi ekstrem, dan hubungan dengan probabilitas pre-test serta post-test. **Sumber:** Diadaptasi dari Ayad *et al*. (2026), Çorbacıoğlu dan Aksel (2023), serta Oehr *et al*. (2025).

7.12 ROC Curve dan AUC

ROC curve adalah grafik yang menggambarkan hubungan antara sensitivitas dan 1-spesifisitas pada berbagai nilai cut-off suatu tes atau model diagnostik (Çorbacıoğlu & Aksel, 2023). ROC curve membantu mengevaluasi kemampuan tes membedakan subjek dengan penyakit dan tanpa penyakit (Çorbacıoğlu & Aksel, 2023). Area under the curve atau AUC merangkum kemampuan diskriminasi tes secara keseluruhan (Çorbacıoğlu & Aksel, 2023). AUC dapat ditafsirkan sebagai probabilitas bahwa subjek sakit memiliki nilai tes lebih tinggi daripada subjek tidak sakit bila arah skor lebih tinggi menunjukkan risiko lebih tinggi (Emir *et al*., 2026).

AUC sebesar 0,5 menunjukkan kemampuan diskriminasi setara tebakan acak, sedangkan AUC mendekati 1,0 menunjukkan diskriminasi yang lebih baik (Çorbacıoğlu & Aksel, 2023). Namun, AUC tidak menunjukkan performa pada satu cut-off spesifik dan tidak selalu mencerminkan kegunaan klinis pada rentang sensitivitas atau spesifisitas tertentu (He *et al*., 2025). Perbandingan AUC antara dua tes harus memperhatikan apakah kedua ROC curve berasal dari subjek yang sama atau berbeda karena korelasi antar AUC dapat memengaruhi inferensi (Emir *et al*., 2026). Kesalahan dalam membandingkan AUC atau menentukan cut-off dapat menyebabkan kesimpulan diagnostik yang keliru (He *et al*., 2025).



Gambar 7. 9. ROC curve dan AUC

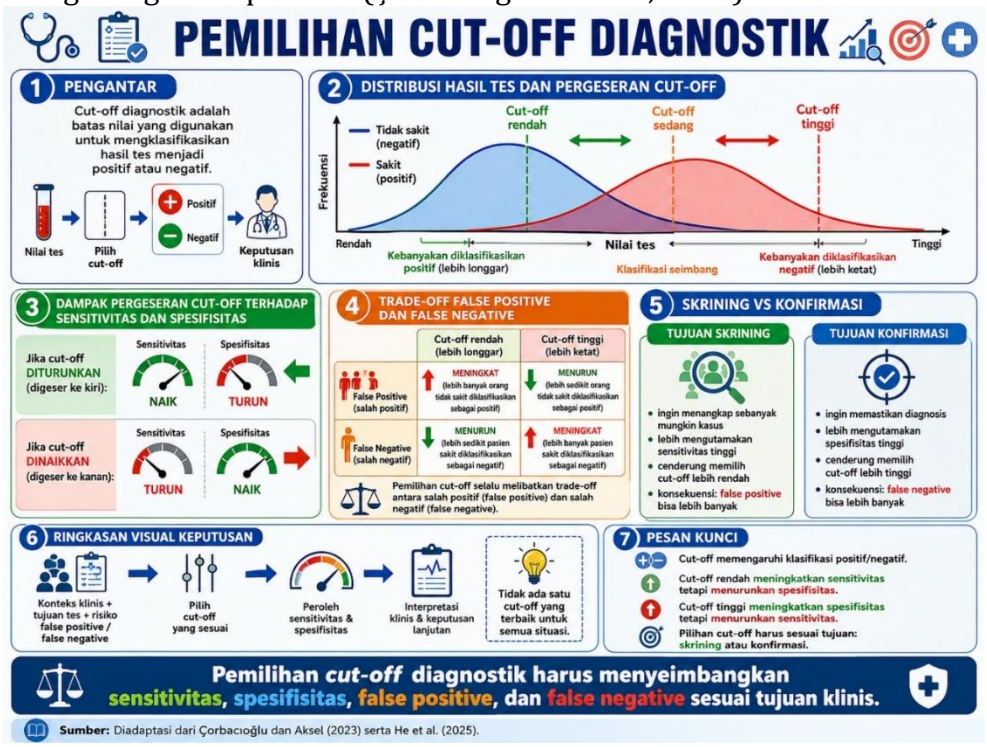
Diagram menunjukkan sumbu sensitivitas, sumbu 1-spesifisitas, beberapa titik cut-off, kurva ROC, AUC, garis tebakan acak, dan interpretasi diskriminasi. **Sumber:** Diadaptasi dari Çorbacıoğlu dan Aksel (2023), He *et al*. (2025), serta Emir *et al*. (2026).

7.13 Pemilihan Cut-Off Diagnostik

Cut-off diagnostik adalah nilai ambang yang digunakan untuk mengklasifikasikan hasil tes sebagai positif atau negatif (Çorbacıoğlu &

Aksel, 2023). Pemilihan cut-off memengaruhi sensitivitas dan spesifisitas secara langsung (Çorbacıoğlu & Aksel, 2023). Jika cut-off diturunkan, sensitivitas sering meningkat tetapi spesifisitas dapat menurun (Çorbacıoğlu & Aksel, 2023). Jika cut-off dinaikkan, spesifisitas sering meningkat tetapi sensitivitas dapat menurun (Çorbacıoğlu & Aksel, 2023).

Cut-off optimal tidak selalu berarti cut-off dengan AUC tertinggi karena AUC mengukur performa global, bukan satu ambang keputusan klinis (He *et al*., 2025). Pemilihan cut-off harus mempertimbangkan konsekuensi false positive dan false negative, prevalensi penyakit, tujuan skrining atau konfirmasi, biaya, keamanan, serta konteks klinis (Çorbacıoğlu & Aksel, 2023). Pada tes skrining penyakit serius, sensitivitas tinggi sering diprioritaskan untuk mengurangi false negative (Çorbacıoğlu & Aksel, 2023). Pada tes konfirmasi yang konsekuensi terapinya berat, spesifisitas tinggi dapat lebih diprioritaskan untuk mengurangi false positive (Çorbacıoğlu & Aksel, 2023).



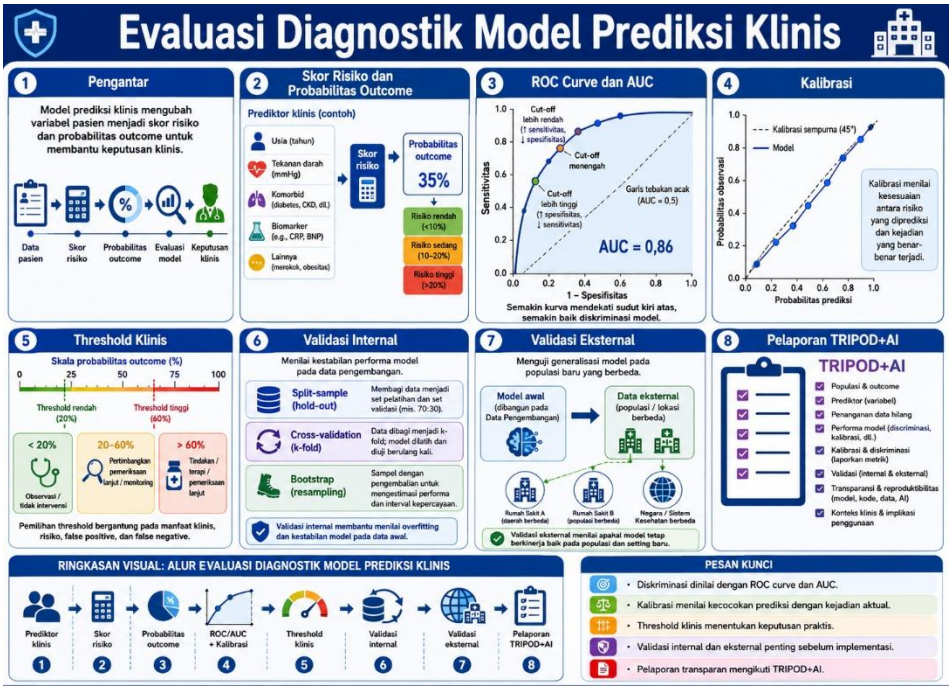
Gambar 7. 10. Pemilihan cut-off diagnostic

Gambar menunjukkan perubahan sensitivitas dan spesifisitas saat cut-off digeser, trade-off false positive dan false negative, serta perbedaan tujuan skrining dan konfirmasi. **Sumber:** Diadaptasi dari Çorbacıoğlu dan Aksel (2023) serta He *et al*. (2025).

7.14 Analisis Diagnostik pada Model Prediksi Klinis

Model prediksi klinis sering menghasilkan skor risiko atau probabilitas outcome, sehingga evaluasinya dapat melibatkan ROC curve, AUC, kalibrasi, dan decision curve analysis (Collins *et al*., 2024). Diskriminasi menunjukkan kemampuan model membedakan subjek yang mengalami outcome dari yang tidak mengalami outcome (Collins *et al*., 2024). Kalibrasi menunjukkan kesesuaian antara probabilitas yang diprediksi model dan kejadian aktual (Collins *et al*., 2024). Model dapat memiliki AUC tinggi tetapi kalibrasi buruk, sehingga evaluasi model tidak boleh hanya berdasarkan ROC curve (Collins *et al*., 2024).

Pedoman TRIPOD+AI menekankan pelaporan transparan untuk model prediksi klinis yang menggunakan regresi maupun machine learning (Collins *et al*., 2024). Studi model prediksi perlu menjelaskan sumber data, outcome, prediktor, missing data, metode pemodelan, validasi, diskriminasi, kalibrasi, dan potensi penerapan klinis (Collins *et al*., 2024). Jika model digunakan untuk keputusan diagnostik, peneliti juga perlu menjelaskan cut-off atau threshold yang digunakan dan konsekuensi klinis dari keputusan tersebut (Collins *et al*., 2024). Dengan demikian, analisis diagnostik pada model prediksi harus menilai performa statistik dan kegunaan klinis secara bersamaan (Collins *et al*., 2024).



Gambar 7. 11. Evaluasi diagnostik model prediksi klinis

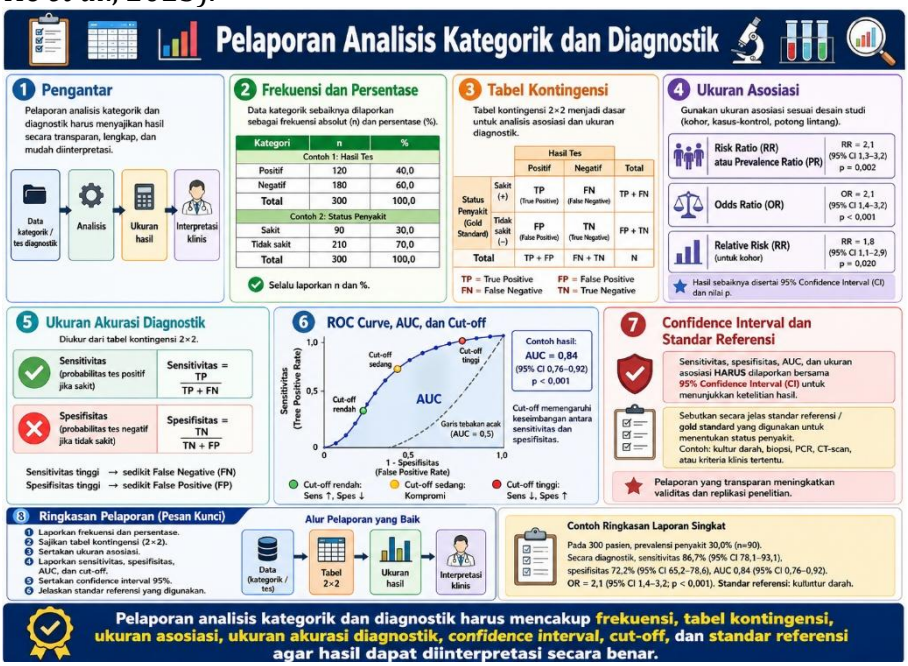
Diagram menunjukkan skor risiko, probabilitas outcome, ROC curve, AUC, kalibrasi, threshold klinis, validasi internal, validasi eksternal, dan pelaporan TRIPOD+AI. **Sumber:** Diadaptasi dari Collins *et al*. (2024), Çorbacıoğlu dan Aksel (2023), serta He *et al*. (2025).

7.15 Pelaporan Analisis Data Kategorik

Pelaporan analisis data kategorik harus menyajikan frekuensi dan persentase pada setiap kategori (Kosin Medical Journal, 2025). Jika chi-square test digunakan, laporan sebaiknya mencantumkan statistik chi-square, derajat kebebasan, nilai p, dan ukuran asosiasi (Kosin Medical Journal, 2025). Jika Fisher's exact test digunakan, laporan perlu menjelaskan alasan penggunaan uji eksak, terutama bila expected count rendah (Chicco *et al*., 2025). Jika McNemar test digunakan, laporan harus menunjukkan bahwa data bersifat berpasangan dan menyajikan pasangan discordant (Sundjaja *et al*., 2023).

Pelaporan analisis diagnostik harus menyertakan tabel 2×2, standar referensi, sensitivitas, spesifisitas, predictive values, likelihood ratios, accuracy, AUC, confidence interval, dan cut-off yang digunakan bila relevan (Ayad *et al*., 2026). Jika ROC curve digunakan, laporan harus menjelaskan apakah AUC digunakan untuk diskriminasi umum atau untuk membandingkan beberapa tes diagnostik (He *et al*., 2025). Jika cut-off dipilih dari data yang sama, peneliti harus menyadari risiko optimisme

dan perlunya validasi eksternal (He *et al*., 2025). Dengan demikian, pelaporan kategorik dan diagnostik harus transparan agar pembaca dapat menilai validitas serta penerapan klinis hasil penelitian (Ayad *et al*., 2026; He *et al*., 2025).



Gambar 7. 12. Pelaporan analisis kategorik dan diagnostic

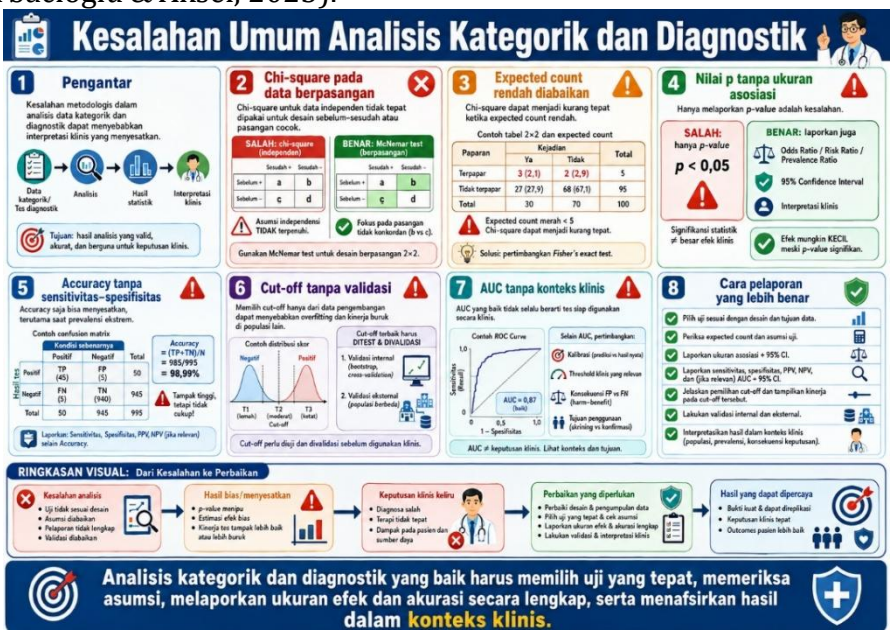
Gambar menunjukkan frekuensi, persentase, tabel kontingensi, ukuran asosiasi, sensitivitas, spesifisitas, AUC, confidence interval, cut-off, dan standar referensi. **Sumber:** Diadaptasi dari Kosin Medical Journal (2025), Sundjaja *et al*. (2023), Ayad *et al*. (2026), dan He *et al*. (2025).

7.16 Kesalahan Umum dalam Analisis Kategorik dan Diagnostik

Kesalahan umum pertama adalah menggunakan chi-square test pada data berpasangan, padahal McNemar test lebih sesuai untuk data nominal dikotom berpasangan (Sundjaja *et al*., 2023). Kesalahan umum kedua adalah menggunakan chi-square test ketika expected count terlalu rendah, padahal Fisher's exact test lebih sesuai pada kondisi tersebut (Chicco *et al*., 2025). Kesalahan umum ketiga adalah hanya melaporkan nilai p tanpa risk ratio, odds ratio, risk difference, atau ukuran asosiasi lain (Kosin Medical Journal, 2025). Kesalahan umum keempat adalah menafsirkan odds ratio sebagai risk ratio tanpa mempertimbangkan prevalensi outcome (Egbuchulem *et al*., 2025).

Kesalahan dalam analisis diagnostik meliputi hanya melaporkan accuracy tanpa sensitivitas dan spesifisitas, mengabaikan pengaruh

prevalensi terhadap nilai prediksi, memilih cut-off tanpa validasi, dan membandingkan AUC dengan metode yang tidak sesuai (Ayad *et al*., 2026; He *et al*., 2025). Accuracy dapat menyesatkan bila prevalensi outcome ekstrem, sehingga harus dilaporkan bersama ukuran diagnostik lain (Oehr *et al*., 2025). AUC juga tidak boleh dianggap sebagai satu-satunya ukuran kegunaan klinis karena AUC tidak menunjukkan konsekuensi false positive dan false negative pada threshold tertentu (He *et al*., 2025). Oleh karena itu, analisis kategorik dan diagnostik harus selalu dihubungkan dengan desain studi, struktur data, dan konteks keputusan klinis (Çorbacıoğlu & Aksel, 2023).

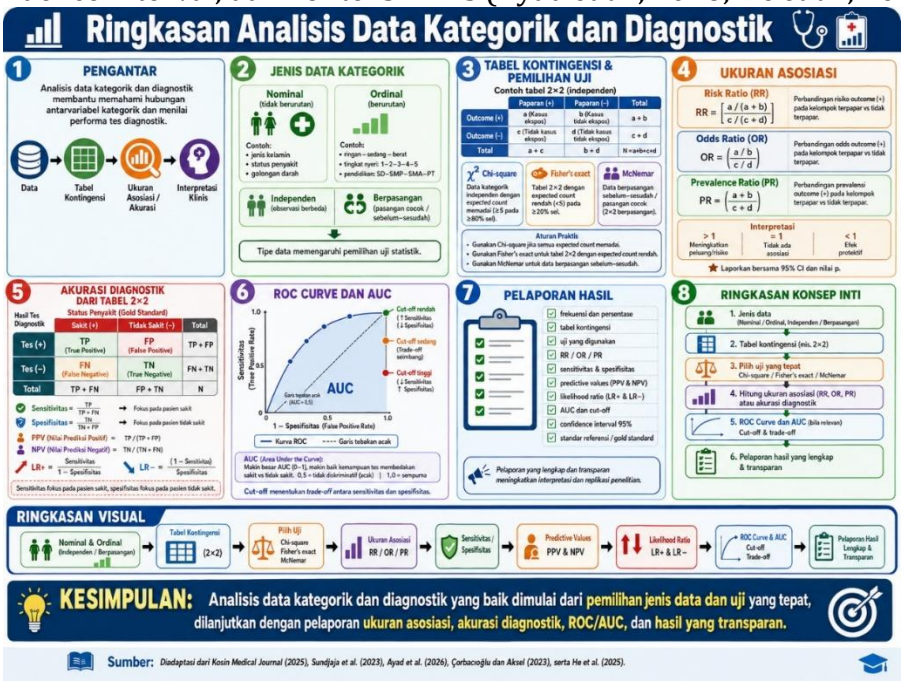


Gambar 7. 13. Kesalahan umum analisis kategorik dan diagnostic
Diagram menunjukkan chi-square pada data berpasangan, expected count rendah, nilai p tanpa ukuran asosiasi, accuracy tanpa sensitivitas-spesifisitas, cut-off tanpa validasi, dan AUC tanpa konteks klinis. **Sumber:** Diadaptasi dari Sundjaja *et al*. (2023), Chicco *et al*. (2025), Ayad *et al*. (2026), He *et al*. (2025), serta Oehr *et al*. (2025).

Bab ini menjelaskan bahwa analisis data kategorik digunakan untuk menilai frekuensi, proporsi, asosiasi, dan perubahan kategori dalam penelitian kesehatan (Kosin Medical Journal, 2025). Chi-square test digunakan untuk dua variabel kategorik independen ketika expected count memadai, sedangkan Fisher's exact test digunakan ketika ukuran sampel kecil atau expected count rendah (Chicco *et al*., 2025). McNemar test digunakan untuk data nominal dikotom berpasangan, terutama pada

desain sebelum-sesudah atau paired diagnostic comparison (Sundjaja *et al*., 2023). Ukuran asosiasi seperti risk difference, risk ratio, dan odds ratio perlu dilaporkan agar interpretasi tidak hanya bergantung pada nilai p (Kosin Medical Journal, 2025).

Bab ini juga menjelaskan bahwa analisis diagnostik menggunakan tabel 2×2 untuk menghitung sensitivitas, spesifisitas, predictive values, likelihood ratios, accuracy, ROC curve, dan AUC (Ayad *et al*., 2026). Sensitivitas dan spesifisitas menggambarkan performa tes terhadap status penyakit, sedangkan predictive values dipengaruhi oleh prevalensi penyakit dalam populasi uji (Ayad *et al*., 2026). ROC curve dan AUC membantu menilai diskriminasi tes atau model, tetapi pemilihan cut-off dan perbandingan AUC harus dilakukan hati-hati agar tidak menghasilkan kesimpulan klinis yang keliru (Çorbacıoğlu & Aksel, 2023; He *et al*., 2025). Dengan demikian, analisis kategorik dan diagnostik yang baik harus memadukan struktur data, ukuran asosiasi, performa diagnostik, confidence interval, dan konteks klinis (Ayad *et al*., 2026; He *et al*., 2025).



Gambar 7. 14. Ringkasan analisis data kategorik dan diagnostic
Diagram merangkum data nominal, ordinal, tabel kontingensi, chi-square, Fisher's exact, McNemar, risk ratio, odds ratio, sensitivitas, spesifisitas, predictive values, likelihood ratio, ROC curve, AUC, dan pelaporan hasil. **Sumber:** Diadaptasi dari Kosin Medical Journal (2025), Sundjaja *et al*. (2023), Ayad *et al*. (2026), Çorbacıoğlu dan Aksel (2023), serta He *et al*. (2025).

Review Questions

1. **Apa yang dimaksud data kategorik?**
   Jawaban: Data kategorik adalah data yang diklasifikasikan ke dalam kelompok atau kategori, seperti status penyakit, jenis kelamin, status paparan, atau kelompok terapi.
2. **Kapan chi-square test digunakan?**
   Jawaban: Chi-square test digunakan untuk menilai hubungan atau perbedaan proporsi antara dua variabel kategorik independen ketika expected count memadai.
3. **Kapan Fisher's exact test lebih tepat digunakan?**
   Jawaban: Fisher's exact test lebih tepat digunakan ketika ukuran sampel kecil atau expected count dalam tabel kontingensi rendah.
4. **Kapan McNemar test digunakan?**
   Jawaban: McNemar test digunakan untuk data nominal dikotom berpasangan, seperti status positif-negatif sebelum dan sesudah intervensi pada subjek yang sama.
5. **Apa perbedaan risk ratio dan odds ratio?**
   Jawaban: Risk ratio membandingkan risiko outcome antar kelompok, sedangkan odds ratio membandingkan odds outcome antar kelompok.
6. **Apa yang dimaksud sensitivitas?**
   Jawaban: Sensitivitas adalah kemampuan tes untuk mengidentifikasi pasien yang benar-benar memiliki penyakit.
7. **Apa yang dimaksud spesifisitas?**
   Jawaban: Spesifisitas adalah kemampuan tes untuk mengidentifikasi individu yang benar-benar tidak memiliki penyakit.
8. **Mengapa nilai prediksi positif dan negatif dipengaruhi prevalensi?**
   Jawaban: Karena probabilitas seseorang benar-benar sakit atau tidak sakit setelah hasil tes bergantung pada seberapa sering penyakit tersebut terdapat dalam populasi yang diuji.
9. **Apa fungsi ROC curve?**
   Jawaban: ROC curve digunakan untuk menilai hubungan sensitivitas dan 1-spesifisitas pada berbagai cut-off serta mengevaluasi kemampuan diskriminatif suatu tes.
10. **Mengapa accuracy tidak boleh digunakan sendirian?**
    Jawaban: Accuracy dapat menyesatkan pada prevalensi penyakit yang sangat rendah atau sangat tinggi, sehingga perlu dilaporkan

bersama sensitivitas, spesifisitas, predictive values, dan likelihood ratios.

DAFTAR PUSTAKA

Ayad, N., *et al*. (2026). EvalTest: A web-based tool for the assessment of diagnostic test performance. *PMC-indexed article*.

Çorbacıoğlu, Ş. K., & Aksel, G. (2023). A guide to interpreting the area under the curve value. *Turkish Journal of Emergency Medicine*.

Egbuchulem, K. I., *et al*. (2025). The odds ratio: A measure of strength in clinical research. *PMC-indexed article*.

Emir, B., *et al*. (2026). Sample size and power analysis for ROC AUC differences in diagnostic studies. *BMC Medical Research Methodology*.

Garrocho-Rangel, A., *et al*. (2024). Fundamentals of nonparametric statistical tests for dental research. *Dentistry Journal, 12*(10), 314.

He, Z., *et al*. (2025). Four overlooked errors in ROC analysis: How to prevent misinterpretation. *BMJ Evidence-Based Medicine, 30*(3), 208–213.

Kaliyadan, F., & Kulkarni, V. (2019). Types of variables, descriptive statistics, and sample size. *Indian Dermatology Online Journal, 10*(1), 82–86.

Kosin Medical Journal. (2025). Common statistical methods used in medical research. *Kosin Medical Journal*.

Oehr, P., *et al*. (2025). Interrelationships among sensitivity, precision, accuracy, predictive values, and ROC curves. *Diagnostics, 15*(4), 410.

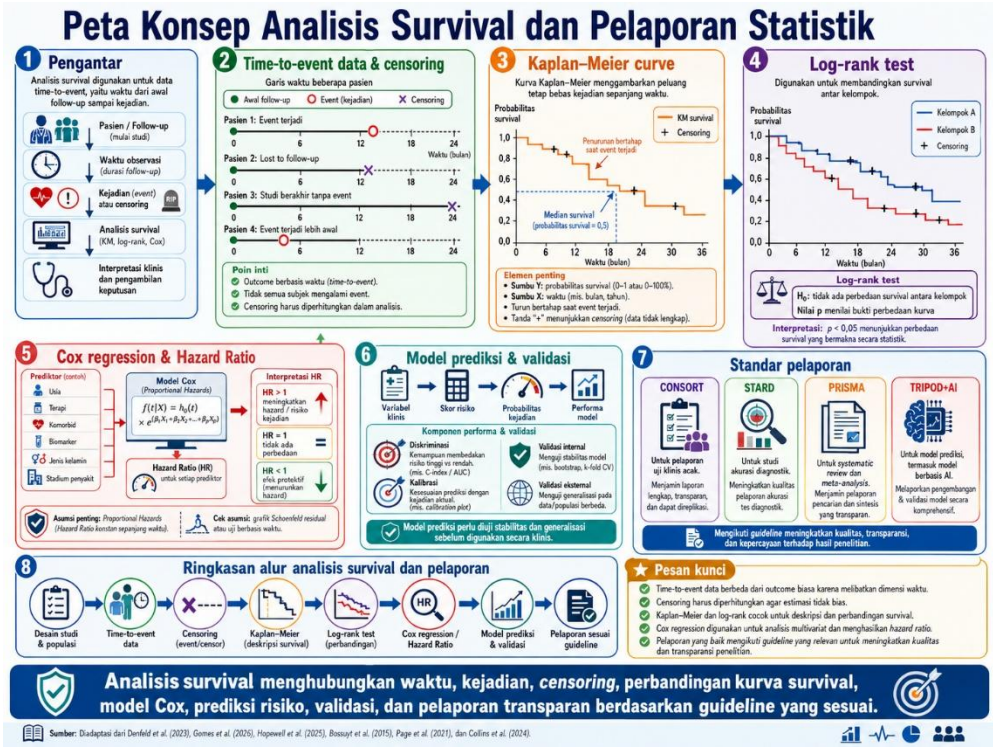
Sundjaja, J. H., Shrestha, R., & Krishan, K. (2023). McNemar and Mann–Whitney U tests. In *StatPearls*. StatPearls Publishing.

BAB 8
ANALISIS SURVIVAL, MODEL PREDIKSI KLINIS, DAN PEDOMAN PELAPORAN STATISTIK

8.1 Pendahuluan

Analisis survival digunakan ketika outcome penelitian berkaitan dengan waktu sampai terjadinya suatu peristiwa, seperti kematian, kekambuhan penyakit, kegagalan terapi, waktu penyembuhan, atau waktu terjadinya komplikasi (Denfeld *et al*., 2023). Analisis survival berbeda dari analisis biasa karena mempertimbangkan **waktu kejadian** dan **status kejadian**, termasuk data tersensor ketika peristiwa belum terjadi sampai akhir pengamatan (Denfeld *et al*., 2023). Dalam penelitian klinis, Kaplan–Meier digunakan untuk memperkirakan probabilitas survival sepanjang waktu, sedangkan Cox proportional hazards regression digunakan untuk menilai hubungan antara prediktor dan hazard setelah mempertimbangkan waktu pengamatan (Denfeld *et al*., 2023; Gomes *et al*., 2026).

Bab ini membahas Kaplan–Meier curve, censoring, log-rank test, Cox regression, hazard ratio, proportional hazards assumption, model prediksi klinis, validasi model, serta pedoman pelaporan seperti CONSORT, STARD, PRISMA, dan TRIPOD+AI. Pedoman pelaporan penting karena kualitas analisis statistik tidak hanya ditentukan oleh uji yang digunakan, tetapi juga oleh kelengkapan laporan metode, asumsi, hasil, ketidakpastian, dan keterbatasan penelitian (Hopewell *et al*., 2025; Collins *et al*., 2024). CONSORT 2025 menyediakan checklist pelaporan randomized trials, sedangkan TRIPOD+AI memberikan panduan pelaporan model prediksi klinis berbasis regresi atau machine learning (Hopewell *et al*., 2025; Collins *et al*., 2024).



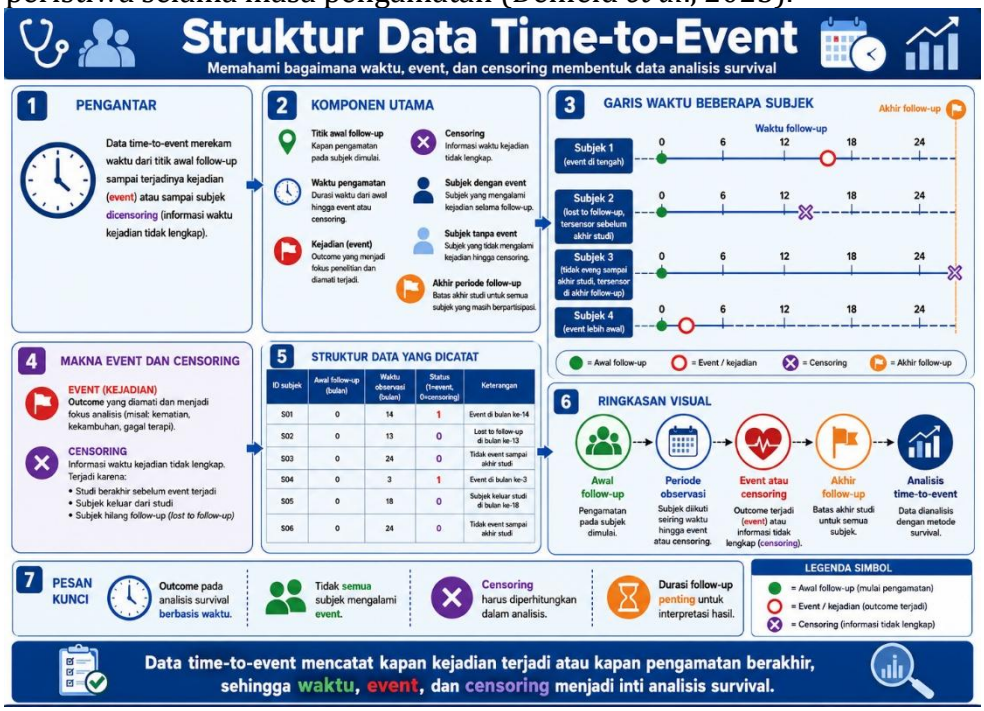
Gambar 8. 1. Peta konsep analisis survival dan pelaporan statistic Ilustrasi menunjukkan time-to-event data, censoring, Kaplan–Meier curve, log-rank test, Cox regression, hazard ratio, model prediksi, validasi, CONSORT, STARD, PRISMA, dan TRIPOD+AI. **Sumber:** Diadaptasi dari Denfeld *et al*. (2023), Gomes *et al*. (2026), Hopewell *et al*. (2025), Bossuyt *et al*. (2015), Page *et al*. (2021), dan Collins *et al*. (2024).

8.2 Konsep Dasar Data Time-to-Event

Data time-to-event adalah data yang mengukur lamanya waktu dari titik awal tertentu sampai terjadinya peristiwa yang telah didefinisikan (Denfeld *et al*., 2023). Titik awal dapat berupa tanggal diagnosis, tanggal terapi, tanggal operasi, tanggal randomisasi, atau tanggal masuk rumah sakit sesuai desain penelitian (Denfeld *et al*., 2023). Peristiwa dapat berupa kematian, relaps, kegagalan implan, komplikasi, kesembuhan, atau outcome klinis lain yang relevan (Denfeld *et al*., 2023). Analisis survival diperlukan karena subjek dapat memiliki waktu follow-up yang berbeda dan sebagian subjek belum mengalami peristiwa saat penelitian berakhir (Denfeld *et al*., 2023).

Dalam analisis survival, setiap subjek biasanya memiliki dua informasi utama, yaitu waktu pengamatan dan status kejadian (Denfeld *et al*., 2023). Status kejadian menunjukkan apakah peristiwa benar-benar

terjadi atau data tersensor (Denfeld *et al*., 2023). Jika pasien masih hidup pada akhir penelitian, hilang follow-up, atau keluar dari studi sebelum peristiwa terjadi, maka data pasien tersebut dapat dikategorikan sebagai tersensor (Denfeld *et al*., 2023). Censoring memungkinkan peneliti tetap menggunakan informasi parsial dari subjek yang belum mengalami peristiwa selama masa pengamatan (Denfeld *et al*., 2023).



Gambar 8. 2. Struktur data time-to-event

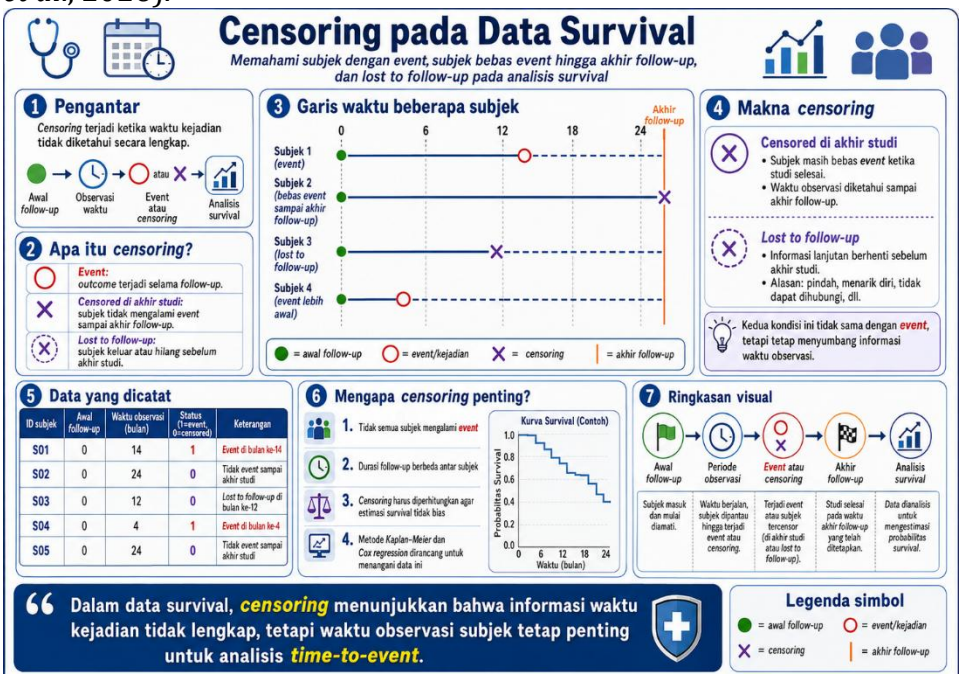
Diagram menunjukkan titik awal follow-up, waktu pengamatan, kejadian, censoring, subjek dengan event, subjek tanpa event, dan akhir periode follow-up. **Sumber:** Diadaptasi dari Denfeld *et al*. (2023) dan Gomes *et al*. (2026).

8.3 Censoring dalam Analisis Survival

Censoring adalah kondisi ketika waktu kejadian sebenarnya tidak diketahui secara lengkap, tetapi diketahui bahwa subjek bebas dari kejadian sampai waktu observasi terakhir (Denfeld *et al*., 2023). Right censoring adalah bentuk censoring yang paling umum dalam penelitian klinis karena kejadian belum terjadi sampai akhir follow-up atau subjek keluar dari penelitian sebelum mengalami kejadian (Denfeld *et al*., 2023). Censoring berbeda dari missing data biasa karena subjek tersensor masih memberikan informasi bahwa peristiwa tidak terjadi sampai waktu tertentu (Denfeld *et al*., 2023). Oleh karena itu, analisis survival dirancang

untuk menggunakan informasi dari subjek dengan event dan subjek tersensor secara bersamaan (Denfeld *et al*., 2023).

Censoring harus dilaporkan secara jelas karena pola censoring dapat memengaruhi interpretasi survival curve dan model hazard (Denfeld *et al*., 2023). Jika censoring terjadi secara acak dan tidak berkaitan dengan risiko kejadian, interpretasi analisis survival lebih dapat diterima (Denfeld *et al*., 2023). Jika censoring berkaitan dengan prognosis, misalnya pasien dengan kondisi lebih buruk lebih sering hilang follow-up, hasil analisis dapat menjadi bias (Denfeld *et al*., 2023). Karena itu, laporan analisis survival perlu menjelaskan jumlah subjek, jumlah event, jumlah censoring, median follow-up, dan alasan censoring bila tersedia (Gomes *et al*., 2026).



Gambar 8. 3. Censoring pada data survival

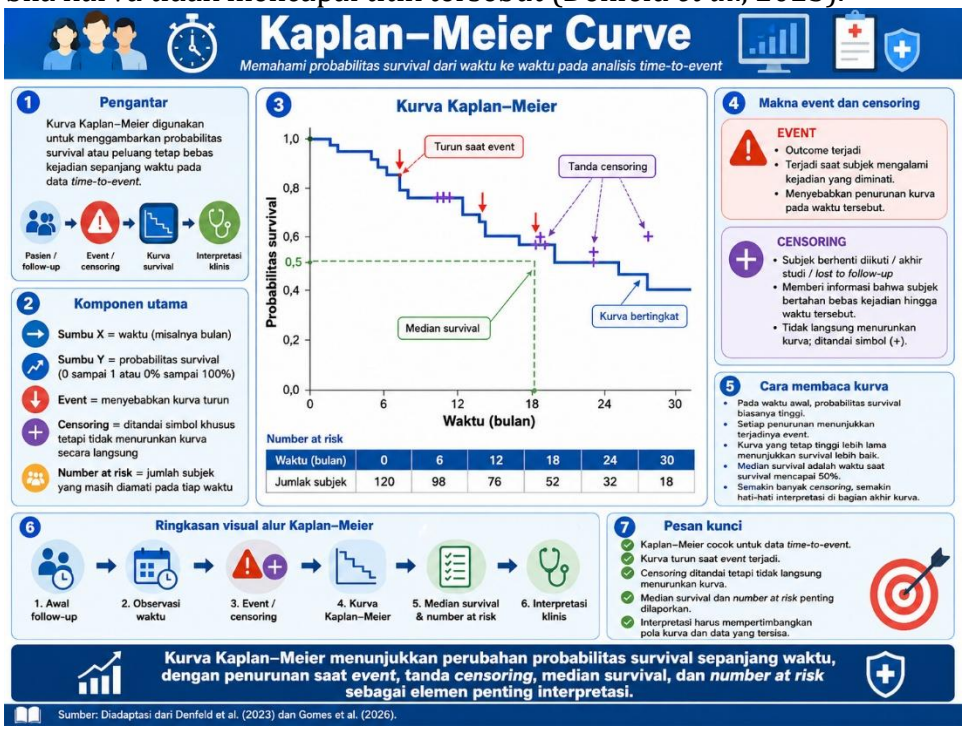
Gambar menunjukkan subjek yang mengalami event, subjek yang masih bebas event sampai akhir follow-up, subjek lost to follow-up, dan tanda censoring pada garis waktu. **Sumber:** Diadaptasi dari Denfeld *et al*. (2023) dan Gomes *et al*. (2026).

8.4 Kaplan–Meier Curve

Kaplan–Meier curve adalah metode nonparametrik untuk memperkirakan probabilitas survival sepanjang waktu (Denfeld *et al*., 2023). Kurva Kaplan–Meier turun setiap kali terjadi event, sedangkan subjek yang tersensor ditandai tetapi tidak menyebabkan penurunan

probabilitas survival pada saat censoring (Denfeld *et al*., 2023). Kaplan–Meier curve dapat digunakan untuk menggambarkan overall survival, progression-free survival, time to recurrence, time to failure, atau outcome time-to-event lain (Gomes *et al*., 2026). Kurva ini sering digunakan karena mampu menampilkan perubahan risiko secara visual selama follow-up (Denfeld *et al*., 2023).

Interpretasi Kaplan–Meier curve perlu memperhatikan jumlah subjek yang masih berisiko pada setiap waktu (Denfeld *et al*., 2023). Jika jumlah subjek yang tersisa sangat sedikit pada akhir follow-up, bagian ekor kurva menjadi kurang stabil (Denfeld *et al*., 2023). Oleh karena itu, Kaplan–Meier plot sebaiknya disertai tabel number at risk di bawah grafik (Gomes *et al*., 2026). Median survival dapat dilaporkan jika probabilitas survival turun sampai 50%, tetapi median survival tidak dapat dihitung bila kurva tidak mencapai titik tersebut (Denfeld *et al*., 2023).



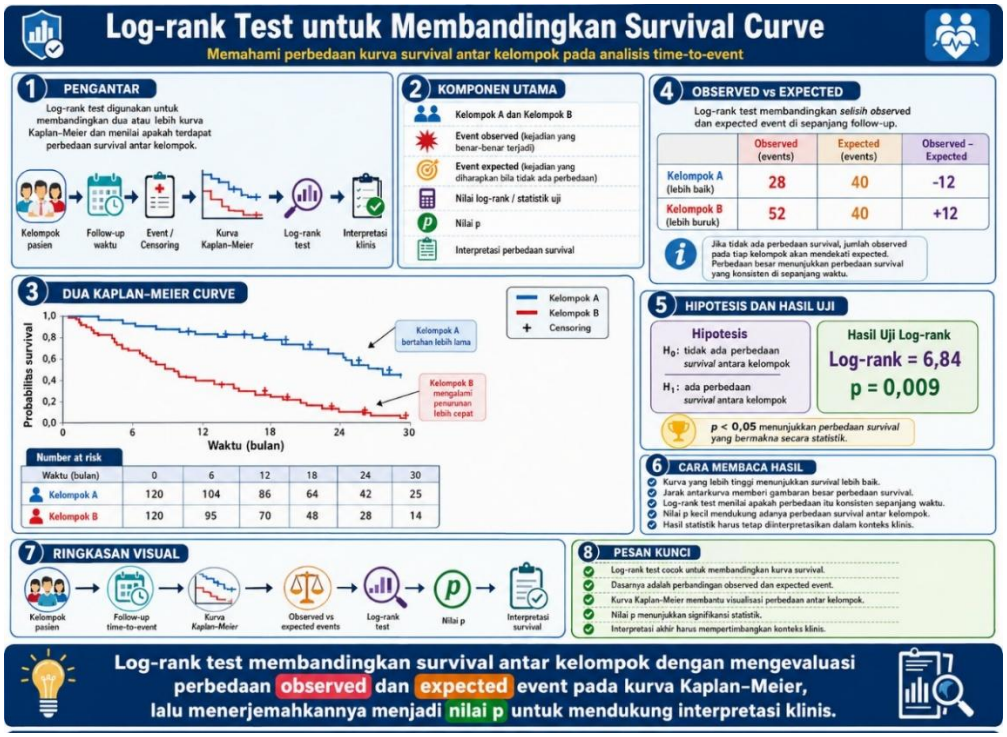
Gambar 8. 4. Kaplan–Meier curve

Diagram menunjukkan kurva bertingkat, penurunan saat event, tanda censoring, median survival, number at risk, dan interpretasi probabilitas survival. **Sumber:** Diadaptasi dari Denfeld *et al*. (2023) dan Gomes *et al*. (2026).

8.5 Log-Rank Test

Log-rank test digunakan untuk membandingkan survival curve antara dua kelompok atau lebih (Denfeld *et al*., 2023). Uji ini membandingkan jumlah event yang diamati dengan jumlah event yang diharapkan pada setiap waktu kejadian berdasarkan hipotesis bahwa tidak ada perbedaan survival antar kelompok (Denfeld *et al*., 2023). Log-rank test sering digunakan untuk membandingkan survival antara kelompok terapi dan kontrol, stadium penyakit, atau kelompok risiko (Gomes *et al*., 2026). Uji ini paling sesuai ketika perbandingan hazard antar kelompok relatif konsisten sepanjang waktu (Denfeld *et al*., 2023).

Log-rank test tidak memberikan ukuran besar efek secara langsung (Denfeld *et al*., 2023). Oleh karena itu, hasil log-rank test sebaiknya dilengkapi dengan estimasi survival, median survival, confidence interval, dan hazard ratio dari model Cox bila sesuai (Gomes *et al*., 2026). Jika survival curve saling berpotongan, log-rank test dapat kurang representatif karena asumsi perbedaan hazard yang konsisten mungkin tidak terpenuhi (Denfeld *et al*., 2023). Dalam kondisi tersebut, peneliti perlu mempertimbangkan metode alternatif atau interpretasi yang lebih hati-hati (Denfeld *et al*., 2023).



Gambar 8. 5. Log-rank test untuk membandingkan survival curve

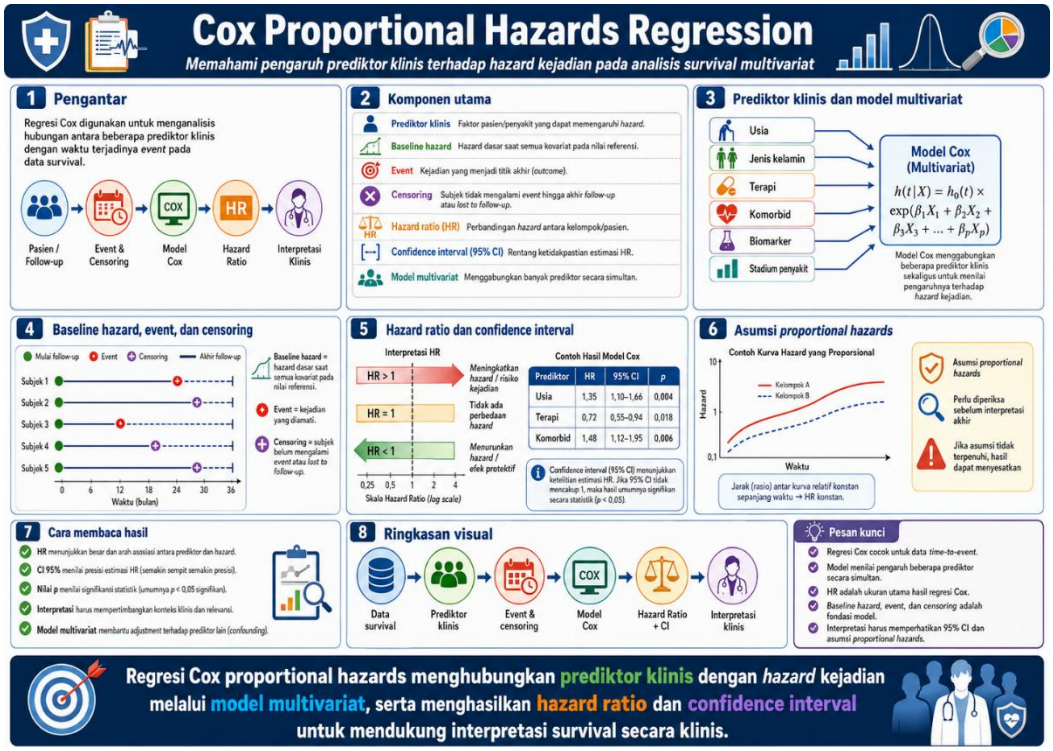
Gambar menunjukkan dua Kaplan–Meier curve, event observed, event expected, nilai log-rank, nilai p, dan interpretasi perbedaan survival antar kelompok. **Sumber:** Diadaptasi dari Denfeld *et al*. (2023) dan Gomes *et al*. (2026).

8.6 Cox Proportional Hazards Regression

Cox proportional hazards regression digunakan untuk menilai hubungan antara satu atau lebih prediktor dengan hazard kejadian dalam data survival (Denfeld *et al*., 2023). Model Cox dapat memasukkan prediktor seperti usia, jenis kelamin, stadium penyakit, terapi, biomarker, atau faktor risiko klinis lain (Gomes *et al*., 2026). Kelebihan model Cox adalah kemampuannya menilai beberapa prediktor secara bersamaan tanpa harus menentukan bentuk baseline hazard secara eksplisit (Denfeld *et al*., 2023). Karena itu, Cox regression sering menjadi metode utama dalam analisis survival multivariat pada penelitian klinis (Gomes *et al*., 2026).

Hasil utama Cox regression biasanya dilaporkan sebagai hazard ratio (Gomes *et al*., 2026). Hazard ratio lebih dari 1 menunjukkan hazard kejadian lebih tinggi pada kelompok atau prediktor tertentu, sedangkan hazard ratio kurang dari 1 menunjukkan hazard kejadian lebih rendah

(Gomes *et al*., 2026). Interpretasi hazard ratio harus dikaitkan dengan periode follow-up, definisi event, dan asumsi proportional hazards (Denfeld *et al*., 2023). Pelaporan Cox regression sebaiknya mencakup hazard ratio, 95% confidence interval, nilai p, jumlah event, variabel dalam model, dan pemeriksaan asumsi (Gomes *et al*., 2026).



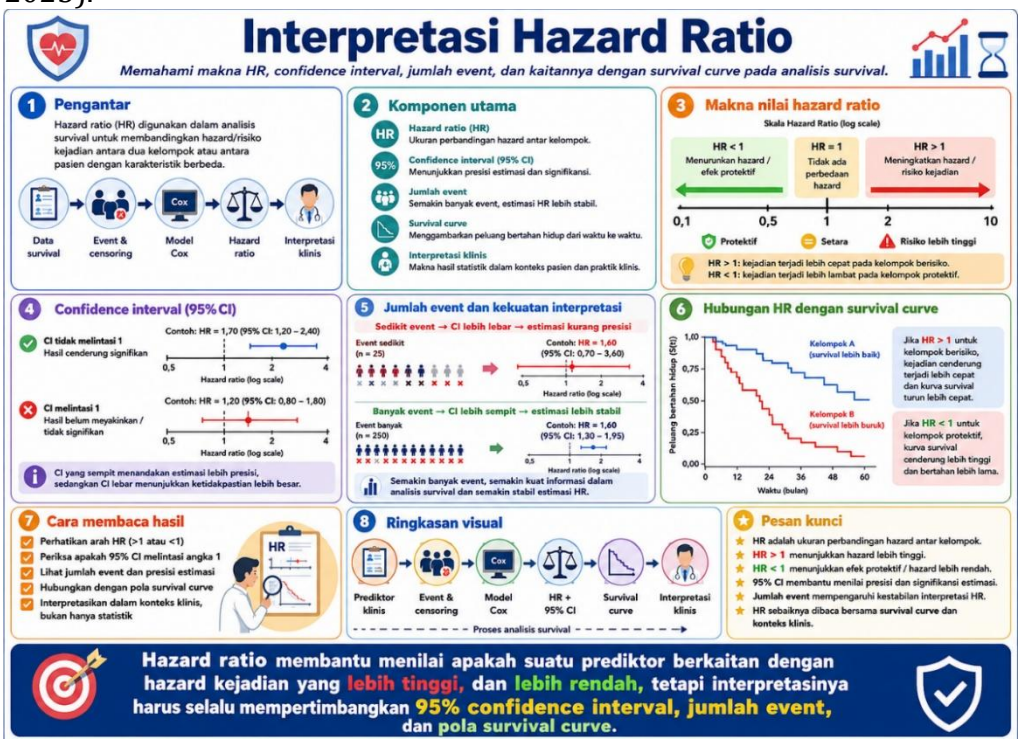
Gambar 8. 6. Cox proportional hazards regression

Diagram menunjukkan prediktor klinis, baseline hazard, hazard ratio, confidence interval, event, censoring, dan model multivariat. **Sumber:** Diadaptasi dari Denfeld *et al*. (2023) dan Gomes *et al*. (2026).

8.7 Hazard Ratio dan Interpretasinya

Hazard ratio adalah perbandingan hazard kejadian antara dua kelompok atau untuk setiap kenaikan satu unit prediktor dalam model Cox (Gomes *et al*., 2026). Hazard menunjukkan laju kejadian sesaat pada waktu tertentu di antara subjek yang masih berisiko (Denfeld *et al*., 2023). Hazard ratio 2,0 tidak berarti risiko kumulatif menjadi dua kali lipat pada semua waktu, tetapi menunjukkan perbandingan hazard berdasarkan model yang digunakan (Denfeld *et al*., 2023). Oleh karena itu, hazard ratio perlu ditafsirkan bersama Kaplan–Meier curve, absolute risk, median survival, dan konteks klinis (Gomes *et al*., 2026).

Hazard ratio harus dilaporkan dengan confidence interval karena estimasi tunggal tidak menunjukkan ketidakpastian (Rovetta *et al*., 2025). Confidence interval yang lebar menunjukkan presisi rendah dan dapat terjadi bila jumlah event sedikit atau variasi data besar (Rovetta *et al*., 2025). Nilai p pada Cox regression tidak boleh digunakan sebagai satu-satunya dasar interpretasi karena nilai p tidak menunjukkan besar efek atau relevansi klinis (Rovetta *et al*., 2025). Interpretasi ilmiah yang baik menggabungkan arah efek, besar efek, interval ketidakpastian, plausibilitas klinis, dan keterbatasan desain penelitian (Rovetta *et al*., 2025).



Gambar 8. 7. Interpretasi hazard ratio

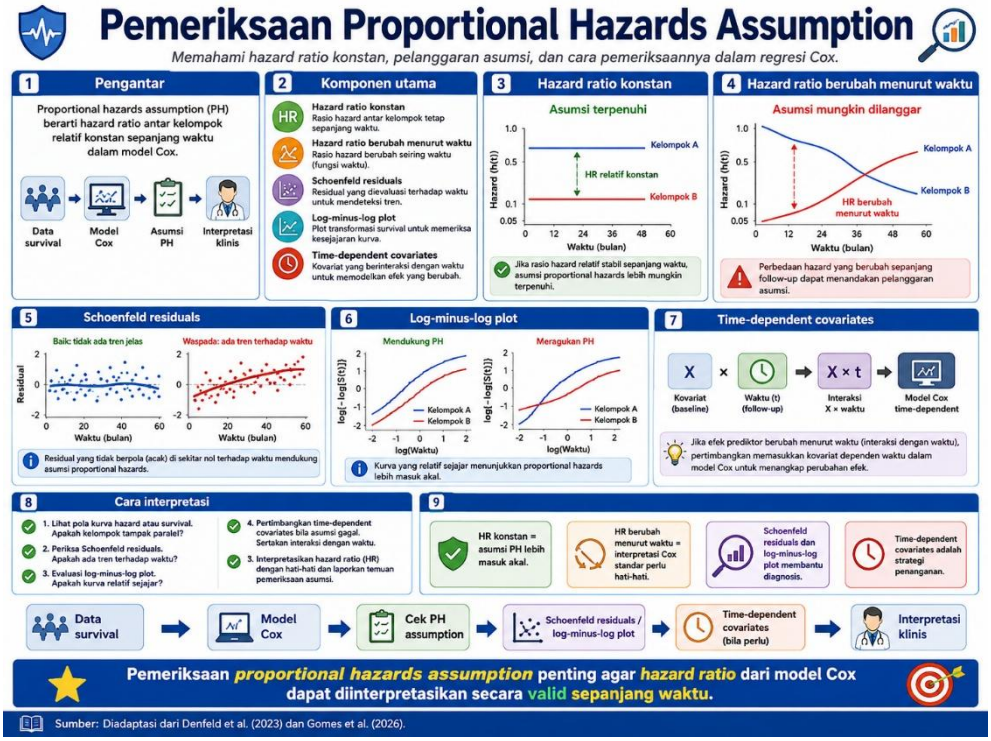
Gambar menunjukkan hazard ratio lebih dari 1, hazard ratio kurang dari 1, confidence interval, jumlah event, dan hubungan hazard ratio dengan survival curve. **Sumber:** Diadaptasi dari Gomes *et al*. (2026), Denfeld *et al*. (2023), dan Rovetta *et al*. (2025).

8.8 Asumsi Proportional Hazards

Asumsi proportional hazards menyatakan bahwa perbandingan hazard antar kelompok relatif konstan sepanjang waktu (Denfeld *et al*., 2023). Jika asumsi ini terpenuhi, hazard ratio dapat ditafsirkan sebagai rasio hazard rata-rata yang relatif stabil selama periode follow-up

(Denfeld *et al*., 2023). Jika asumsi ini dilanggar, hazard ratio tunggal dapat menyederhanakan hubungan yang sebenarnya berubah menurut waktu (Gomes *et al*., 2026). Oleh karena itu, pemeriksaan asumsi proportional hazards merupakan bagian penting dari Cox regression (Gomes *et al*., 2026).

Asumsi proportional hazards dapat diperiksa melalui grafik log-minus-log, Schoenfeld residuals, atau interaksi prediktor dengan waktu (Denfeld *et al*., 2023). Jika asumsi tidak terpenuhi, peneliti dapat mempertimbangkan stratified Cox model, time-dependent covariates, analisis berdasarkan periode waktu, atau model survival alternatif (Gomes *et al*., 2026). Pelaporan Cox regression yang baik seharusnya menyebutkan apakah asumsi proportional hazards telah diperiksa (Gomes *et al*., 2026). Tanpa pemeriksaan ini, pembaca sulit menilai validitas interpretasi hazard ratio (Gomes *et al*., 2026).



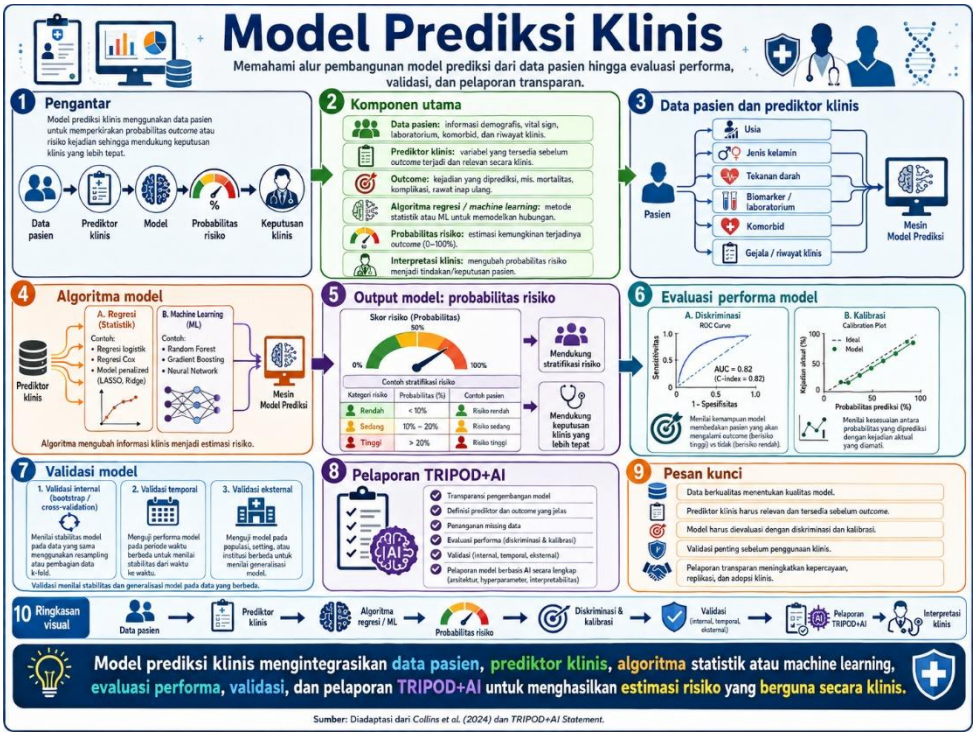
Gambar 8. 8. Pemeriksaan proportional hazards assumption

Diagram menunjukkan hazard ratio konstan, hazard ratio berubah menurut waktu, Schoenfeld residuals, log-minus-log plot, dan time-dependent covariates. **Sumber:** Diadaptasi dari Denfeld *et al*. (2023) dan Gomes *et al*. (2026).

8.9 Model Prediksi Klinis

Model prediksi klinis digunakan untuk memperkirakan probabilitas outcome pada individu berdasarkan kombinasi prediktor klinis, laboratorium, radiologis, genetik, atau digital (Collins *et al*., 2024). Model prediksi dapat dikembangkan menggunakan regresi, machine learning, atau kombinasi metode statistik lain (Collins *et al*., 2024). Outcome model prediksi dapat berupa penyakit, kematian, komplikasi, respons terapi, atau risiko kejadian dalam jangka waktu tertentu (Collins *et al*., 2024). Model prediksi harus dievaluasi berdasarkan diskriminasi, kalibrasi, validasi, dan kegunaan klinis sebelum diterapkan dalam praktik (Collins *et al*., 2024).

TRIPOD+AI memberikan panduan pelaporan untuk studi model prediksi klinis yang menggunakan regresi maupun machine learning (Collins *et al*., 2024). Pedoman ini menggantikan TRIPOD 2015 dan memperluas checklist agar lebih sesuai dengan perkembangan model prediksi modern termasuk AI (Collins *et al*., 2024). Laporan model prediksi perlu menjelaskan sumber data, kriteria peserta, outcome, prediktor, penanganan missing data, metode pemodelan, validasi, diskriminasi, kalibrasi, dan aksesibilitas model (Collins *et al*., 2024). Transparansi pelaporan diperlukan agar pembaca dapat menilai risiko bias, reproducibility, dan kesiapan penerapan klinis model (Collins *et al*., 2024).



Gambar 8. 9. Model prediksi klinis

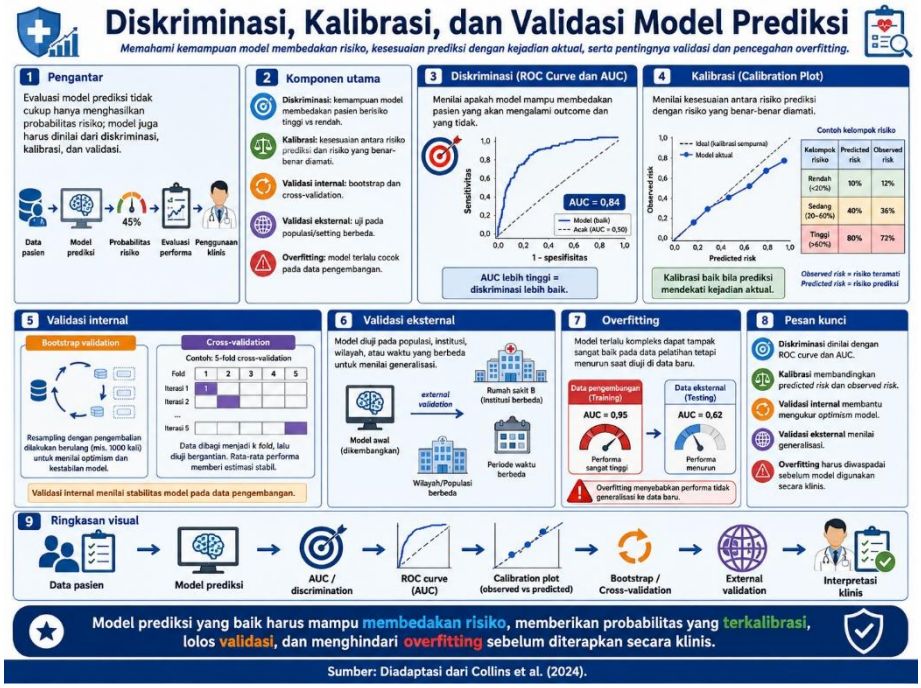
Gambar menunjukkan data pasien, prediktor klinis, algoritma regresi atau machine learning, probabilitas risiko, diskriminasi, kalibrasi, validasi, dan pelaporan TRIPOD+AI. **Sumber:** Diadaptasi dari Collins *et al*. (2024) dan TRIPOD+AI Statement.

8.10 Diskriminasi, Kalibrasi, dan Validasi Model

Diskriminasi adalah kemampuan model membedakan subjek yang mengalami outcome dari subjek yang tidak mengalami outcome (Collins *et al*., 2024). AUC atau C-statistic sering digunakan untuk menilai diskriminasi pada model outcome biner atau time-to-event tertentu (Collins *et al*., 2024). Kalibrasi adalah kesesuaian antara probabilitas yang diprediksi model dan kejadian aktual dalam data (Collins *et al*., 2024). Model dapat memiliki diskriminasi tinggi tetapi kalibrasi buruk, sehingga kedua aspek perlu dilaporkan (Collins *et al*., 2024).

Validasi internal digunakan untuk menilai performa model pada data pengembangan dengan teknik seperti bootstrap atau cross-validation (Collins *et al*., 2024). Validasi eksternal menguji model pada dataset yang berbeda dari data pengembangan dan lebih kuat untuk menilai generalisasi model (Collins *et al*., 2024). Overfitting terjadi ketika model terlalu menyesuaikan diri dengan data pengembangan sehingga

performanya menurun pada data baru (Collins *et al*., 2024). Oleh karena itu, model prediksi yang baik harus dilaporkan dengan performa optimism-adjusted dan validasi eksternal bila memungkinkan (Collins *et al*., 2024).

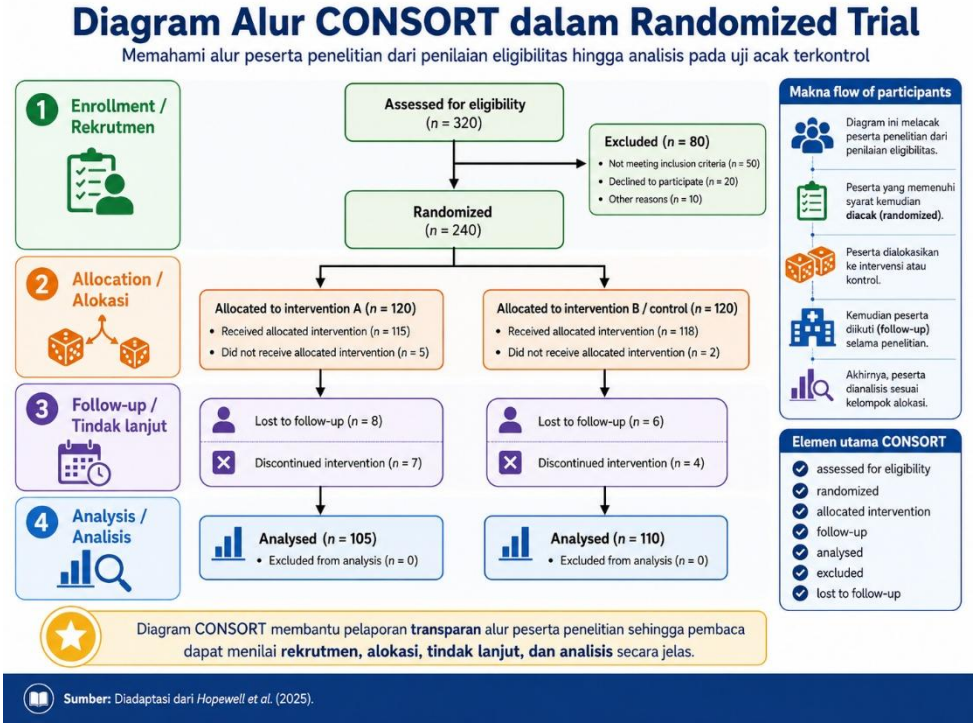


Gambar 8. 10. Diskriminasi, kalibrasi, dan validasi model prediksi
Diagram menunjukkan AUC, calibration plot, observed risk, predicted risk, bootstrap validation, cross-validation, external validation, dan overfitting. **Sumber:** Diadaptasi dari Collins *et al*. (2024).

8.11 CONSORT untuk Randomized Trials

CONSORT adalah pedoman pelaporan randomized trials yang bertujuan meningkatkan kelengkapan dan transparansi laporan penelitian uji acak (Hopewell *et al*., 2025). CONSORT 2025 terdiri atas checklist 30 item dan diagram alur peserta untuk membantu penulis melaporkan desain, metode, hasil, dan interpretasi randomized trials secara lengkap (Hopewell *et al*., 2025). CONSORT penting karena randomized trials sering digunakan untuk menilai manfaat dan risiko intervensi kesehatan (Hopewell *et al*., 2025). Pelaporan yang tidak lengkap dapat menghambat pembaca dalam menilai risiko bias, validitas internal, dan penerapan hasil penelitian (Hopewell *et al*., 2025).

CONSORT 2025 menambahkan dan merevisi beberapa item, termasuk struktur baru yang menekankan aspek open science (Hopewell *et al*., 2025). Diagram alur CONSORT membantu menampilkan jumlah peserta yang dinilai kelayakannya, dirandomisasi, menerima intervensi, menyelesaikan follow-up, dan dianalisis (Hopewell *et al*., 2025). Dalam konteks statistik, CONSORT mendukung pelaporan outcome primer, metode randomisasi, blinding, ukuran sampel, analisis utama, effect size, confidence interval, dan harm (Hopewell *et al*., 2025). Oleh karena itu, peneliti yang melakukan uji klinis acak perlu menggunakan CONSORT sejak tahap penulisan protokol dan artikel (Hopewell *et al*., 2025).



Gambar 8. 11. Diagram alur CONSORT dalam randomized trial

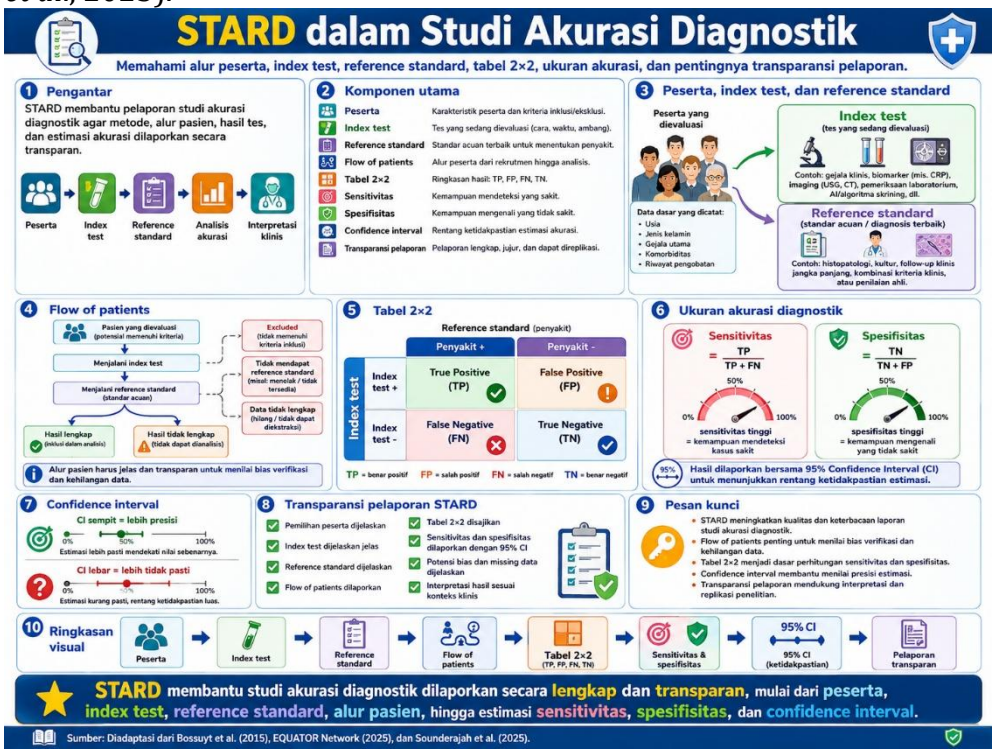
Gambar menunjukkan assessed for eligibility, randomized, allocated intervention, follow-up, analysed, excluded, lost to follow-up, dan flow of participants. **Sumber:** Diadaptasi dari Hopewell *et al*. (2025).

8.12 STARD untuk Studi Akurasi Diagnostik

STARD adalah pedoman pelaporan studi akurasi diagnostik yang bertujuan meningkatkan kelengkapan laporan penelitian diagnostik (Bossuyt *et al*., 2015). STARD 2015 memuat daftar 30 item penting yang seharusnya dilaporkan dalam studi diagnostic accuracy (Bossuyt *et al*., 2015). Pedoman ini membantu pembaca menilai populasi studi, index

test, reference standard, flow of participants, estimasi akurasi, dan risiko bias (Bossuyt *et al*., 2015). EQUATOR Network mencatat bahwa STARD 2015 disediakan sebagai reporting guideline untuk studi akurasi diagnostik (EQUATOR Network, 2025).

STARD penting karena studi diagnostik dapat menghasilkan interpretasi keliru bila index test, reference standard, cut-off, atau flow pasien tidak dilaporkan dengan jelas (Bossuyt *et al*., 2015). Dalam studi diagnostik berbasis AI, STARD-AI dikembangkan untuk menangani kebutuhan pelaporan khusus sistem kecerdasan buatan dalam diagnostic test accuracy studies (Sounderajah *et al*., 2025). STARD-AI menyediakan kriteria minimum agar studi diagnostik AI dilaporkan secara lengkap dan transparan (Sounderajah *et al*., 2025). Oleh karena itu, studi diagnostik konvensional dapat merujuk STARD 2015, sedangkan studi diagnostik AI sebaiknya memperhatikan STARD-AI (Bossuyt *et al*., 2015; Sounderajah *et al*., 2025).



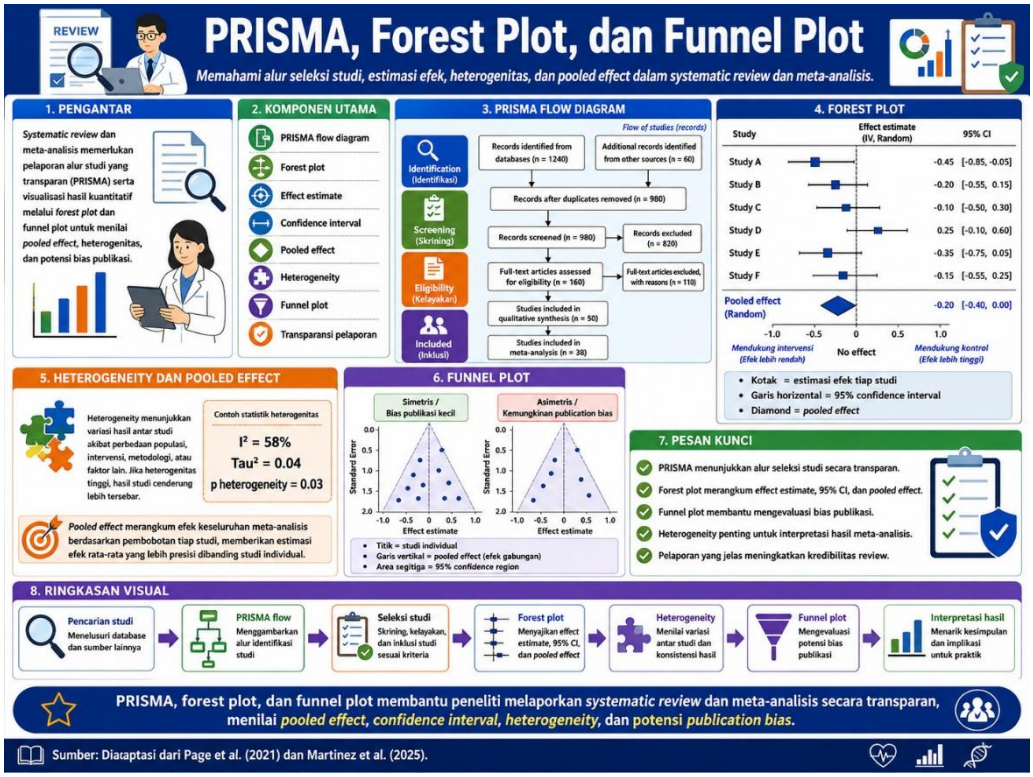
Gambar 8. 12. STARD dalam studi akurasi diagnostic

Diagram menunjukkan peserta, index test, reference standard, flow of patients, tabel 2×2, sensitivitas, spesifisitas, confidence interval, dan transparansi pelaporan.**Sumber:** Diadaptasi dari Bossuyt *et al*. (2015), EQUATOR Network (2025), dan Sounderajah *et al*. (2025).

8.13 PRISMA untuk Systematic Review dan Meta-Analysis

PRISMA 2020 adalah pedoman pelaporan systematic review dan meta-analysis yang membantu penulis melaporkan alasan review dilakukan, metode yang digunakan, dan hasil yang ditemukan secara lengkap (Page *et al*., 2021). PRISMA terutama digunakan untuk review yang mengevaluasi efek intervensi, tetapi itemnya juga sering diadaptasi untuk berbagai jenis systematic review (PRISMA Statement, 2025). PRISMA flow diagram membantu melaporkan proses identifikasi, screening, eligibility, exclusion, dan inclusion studi (Page *et al*., 2021). Pelaporan PRISMA yang baik meningkatkan transparansi dan reproducibility systematic review (Page *et al*., 2021).

Dalam meta-analysis, forest plot adalah grafik utama untuk menampilkan estimasi efek dan confidence interval dari setiap studi serta efek gabungan (Page *et al*., 2021). Funnel plot sering digunakan untuk mengevaluasi kemungkinan small-study effects atau publication bias, meskipun interpretasinya harus hati-hati karena asimetri dapat dipengaruhi oleh banyak faktor (Martinez *et al*., 2025). Meta-analysis harus melaporkan model fixed effect atau random effects, heterogeneity, I^2, analisis sensitivitas, dan risk of bias agar pembaca dapat menilai kekuatan bukti (Martinez *et al*., 2025). Oleh karena itu, systematic review dan meta-analysis membutuhkan pelaporan metodologis yang lebih rinci daripada ringkasan naratif biasa (Page *et al*., 2021; Martinez *et al*., 2025).



Gambar 8. 13. PRISMA, forest plot, dan funnel plot

Gambar menunjukkan PRISMA flow diagram, forest plot dengan effect estimate dan confidence interval, funnel plot, heterogeneity, dan pooled effect. **Sumber:** Diadaptasi dari Page *et al*. (2021) dan Martinez *et al*. (2025).

8.14 Pelaporan Nilai p, Confidence Interval, dan Effect Size

Nilai p menunjukkan kompatibilitas data dengan model statistik dan hipotesis nol, tetapi tidak menunjukkan kebenaran hipotesis atau besar efek secara langsung (Rovetta *et al*., 2025). Confidence interval menunjukkan rentang nilai parameter yang kompatibel dengan data dan model pada tingkat kepercayaan tertentu (Rovetta *et al*., 2025). Effect size menunjukkan besar efek, perbedaan, hubungan, atau asosiasi yang lebih bermakna untuk interpretasi ilmiah daripada nilai p saja (Rovetta *et al*., 2025). Pelaporan yang baik harus menggabungkan effect size, confidence interval, nilai p, dan konteks klinis (Rovetta *et al*., 2025).

Interpretasi "signifikan" dan "tidak signifikan" secara dikotomis dapat menyesatkan karena hasil p > 0,05 tidak selalu membuktikan tidak ada efek (Rovetta *et al*., 2025). Confidence interval yang lebar menunjukkan ketidakpastian tinggi dan dapat mencakup efek yang

bermakna secara klinis meskipun nilai p tidak signifikan (Rovetta *et al*., 2025). Sebaliknya, nilai p sangat kecil dapat muncul pada sampel besar meskipun effect size kecil dan kurang bermakna secara klinis (Rovetta *et al*., 2025). Oleh karena itu, peneliti perlu menafsirkan hasil statistik secara kuantitatif, bukan hanya berdasarkan batas p < 0,05 (Rovetta *et al*., 2025).



Gambar 8. 14. Pelaporan hasil statistik yang lengkap

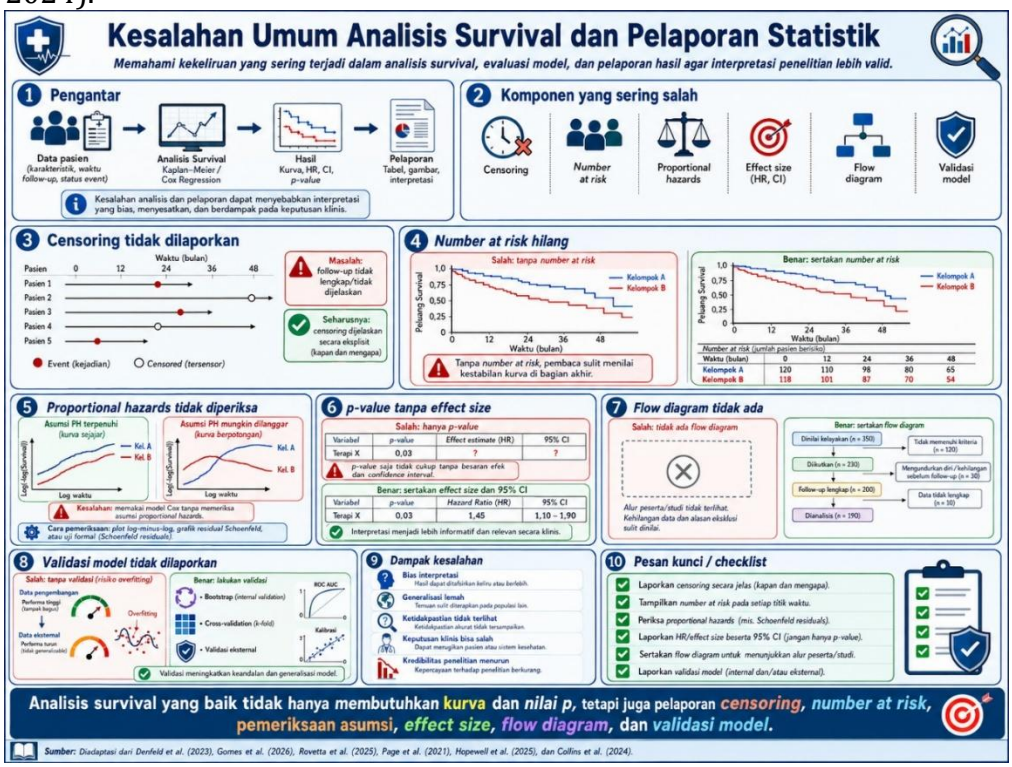
Diagram menunjukkan estimasi efek, confidence interval, nilai p, effect size, relevansi klinis, ketidakpastian, dan interpretasi non-dikotomis. **Sumber:** Diadaptasi dari Rovetta *et al*. (2025).

8.15 Kesalahan Umum dalam Analisis Survival dan Pelaporan Statistik

Kesalahan umum dalam analisis survival adalah mengabaikan censoring, tidak melaporkan number at risk, menafsirkan bagian ekor Kaplan–Meier curve yang hanya memiliki sedikit subjek, dan tidak memeriksa proportional hazards assumption pada model Cox (Denfeld *et al*., 2023; Gomes *et al*., 2026). Kesalahan lain adalah menafsirkan hazard ratio sebagai risk ratio tanpa menjelaskan konteks waktu dan hazard (Denfeld *et al*., 2023). Peneliti juga dapat keliru menyimpulkan perbedaan survival hanya berdasarkan log-rank p-value tanpa melaporkan median survival atau hazard ratio (Gomes *et al*., 2026). Oleh karena itu, analisis

survival harus dilaporkan dengan grafik, jumlah event, censoring, number at risk, estimasi efek, dan asumsi model (Gomes *et al*., 2026).

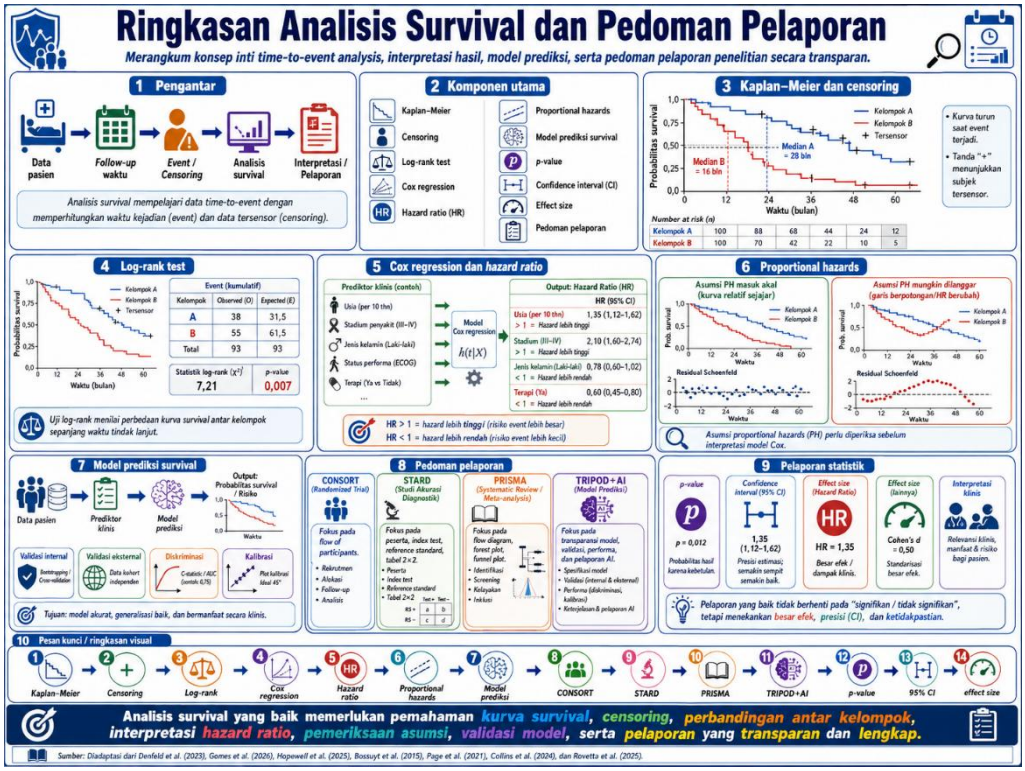
Kesalahan umum dalam pelaporan statistik adalah menyajikan nilai p tanpa effect size dan confidence interval (Rovetta *et al*., 2025). Kesalahan lain adalah tidak menyebutkan pedoman pelaporan yang digunakan, tidak menjelaskan metode penanganan missing data, tidak menyajikan flow diagram, dan tidak melaporkan validasi model prediksi (Hopewell *et al*., 2025; Collins *et al*., 2024). Dalam systematic review, kesalahan pelaporan dapat mencakup tidak menyajikan PRISMA flow diagram, tidak menjelaskan risk of bias, dan tidak melaporkan heterogeneity (Page *et al*., 2021). Dengan demikian, pelaporan statistik harus dirancang untuk membuat proses penelitian dapat dinilai, direplikasi, dan diterapkan secara kritis (Page *et al*., 2021; Collins *et al*., 2024).



Gambar 8. 15. Kesalahan umum analisis survival dan pelaporan statistic
Gambar menunjukkan censoring tidak dilaporkan, number at risk hilang, proportional hazards tidak diperiksa, p-value tanpa effect size, flow diagram tidak ada, dan validasi model tidak dilaporkan. **Sumber:** Diadaptasi dari Denfeld *et al*. (2023), Gomes *et al*. (2026), Rovetta *et al*. (2025), Page *et al*. (2021), Hopewell *et al*. (2025), dan Collins *et al*. (2024).

Bab ini menjelaskan bahwa analisis survival digunakan untuk data time-to-event yang melibatkan waktu kejadian, status event, dan censoring (Denfeld *et al*., 2023). Kaplan–Meier curve digunakan untuk memperkirakan survival probability, log-rank test digunakan untuk membandingkan survival curve, dan Cox proportional hazards regression digunakan untuk menilai hubungan prediktor dengan hazard kejadian (Denfeld *et al*., 2023; Gomes *et al*., 2026). Hazard ratio harus ditafsirkan bersama confidence interval, asumsi proportional hazards, number at risk, dan konteks klinis (Gomes *et al*., 2026). Analisis survival yang baik harus melaporkan event, censoring, follow-up, survival estimates, hazard ratio, dan pemeriksaan asumsi model (Denfeld *et al*., 2023; Gomes *et al*., 2026).

Bab ini juga menegaskan bahwa pelaporan statistik harus mengikuti pedoman yang sesuai dengan desain studi (Hopewell *et al*., 2025; Bossuyt *et al*., 2015; Page *et al*., 2021; Collins *et al*., 2024). CONSORT digunakan untuk randomized trials, STARD digunakan untuk diagnostic accuracy studies, PRISMA digunakan untuk systematic reviews dan meta-analyses, serta TRIPOD+AI digunakan untuk model prediksi klinis berbasis regresi atau machine learning (Hopewell *et al*., 2025; Bossuyt *et al*., 2015; Page *et al*., 2021; Collins *et al*., 2024). Nilai p perlu dilaporkan bersama effect size dan confidence interval agar interpretasi tidak jatuh pada dikotomi “signifikan” versus “tidak signifikan” (Rovetta *et al*., 2025). Dengan demikian, kualitas statistik bukan hanya ditentukan oleh uji yang dipilih, tetapi juga oleh transparansi desain, pelaporan, asumsi, ketidakpastian, dan interpretasi klinis (Rovetta *et al*., 2025; Collins *et al*., 2024).



Gambar 8. 16. Ringkasan analisis survival dan pedoman pelaporan

Diagram merangkum Kaplan–Meier, censoring, log-rank, Cox regression, hazard ratio, proportional hazards, model prediksi, CONSORT, STARD, PRISMA, TRIPOD+AI, p-value, confidence interval, dan effect size. **Sumber:** Diadaptasi dari Denfeld *et al*. (2023), Gomes *et al*. (2026), Hopewell *et al*. (2025), Bossuyt *et al*. (2015), Page *et al*. (2021), Collins *et al*. (2024), dan Rovetta *et al*. (2025).

Review Questions

1. **Apa yang dimaksud analisis survival?**
   Jawaban: Analisis survival adalah analisis statistik untuk data waktu sampai terjadinya suatu peristiwa, seperti kematian, relaps, komplikasi, atau kegagalan terapi.
2. **Apa yang dimaksud censoring?**
   Jawaban: Censoring adalah kondisi ketika subjek belum mengalami event sampai akhir follow-up atau keluar dari pengamatan sebelum event terjadi.
3. **Apa fungsi Kaplan–Meier curve?**
   Jawaban: Kaplan–Meier curve digunakan untuk memperkirakan dan menampilkan probabilitas survival sepanjang waktu.
4. **Kapan log-rank test digunakan?**
   Jawaban: Log-rank test digunakan untuk membandingkan survival curve antara dua kelompok atau lebih.
5. **Apa fungsi Cox proportional hazards regression?**
   Jawaban: Cox regression digunakan untuk menilai hubungan antara prediktor dan hazard kejadian pada data survival.
6. **Apa yang dimaksud hazard ratio?**
   Jawaban: Hazard ratio adalah perbandingan hazard kejadian antara dua kelompok atau untuk setiap kenaikan satu unit prediktor dalam model Cox.
7. **Mengapa proportional hazards assumption penting?**
   Jawaban: Asumsi ini penting karena hazard ratio dalam Cox regression ditafsirkan sebagai rasio hazard yang relatif konstan sepanjang waktu.
8. **Apa fungsi TRIPOD+AI?**
   Jawaban: TRIPOD+AI memberikan panduan pelaporan model prediksi klinis yang menggunakan regresi atau machine learning.
9. **Apa perbedaan CONSORT, STARD, dan PRISMA?**
   Jawaban: CONSORT digunakan untuk randomized trials, STARD untuk studi akurasi diagnostik, dan PRISMA untuk systematic review serta meta-analysis.
10. **Mengapa nilai p tidak cukup dilaporkan sendirian?**
    Jawaban: Nilai p tidak menunjukkan besar efek atau relevansi klinis, sehingga perlu dilaporkan bersama effect size dan confidence interval.

DAFTAR PUSTAKA

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L., Lijmer, J. G., Moher, D., Rennie, D., de Vet, H. C. W., Kressel, H. Y., Rifai, N., Golub, R. M., Altman, D. G., Hooft, L., Korevaar, D. A., & Cohen, J. F. (2015). STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *BMJ, 351*, h5527.

Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., Ghassemi, M., Liu, X., Reitsma, J. B., van Smeden, M., Boulesteix, A. L., Camaradou, J. C., Celi, L. A., Denaxas, S., Denniston, A. K., Glocker, B., Golub, R. M., Harvey, H., Heinze, G., *et al*. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ, 385*, e078378.

Denfeld, Q. E., *et al*. (2023). Survival analysis 101: An easy start guide to analyzing time-to-event data. *PMC-indexed article*.

Gomes, A., *et al*. (2026). Cox regression in survival analysis: Practical insights for clinicians. *Acta Médica Portuguesa*.

Hopewell, S., Chan, A.-W., Collins, G. S., Hróbjartsson, A., Moher, D., Schulz, K. F., *et al*. (2025). CONSORT 2025 statement: Updated guideline for reporting randomised trials. *The Lancet*.

Martinez, E. C., *et al*. (2025). A comprehensive guide to conduct a systematic review and meta-analysis. *PMC-indexed article*.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., *et al*. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ, 372*, n160.

Rovetta, A., Piretta, L., & Mansournia, M. A. (2025). p-values and confidence intervals as compatibility measures: Guidelines for interpreting statistical studies in clinical research. *The Lancet Regional Health – Southeast Asia*.

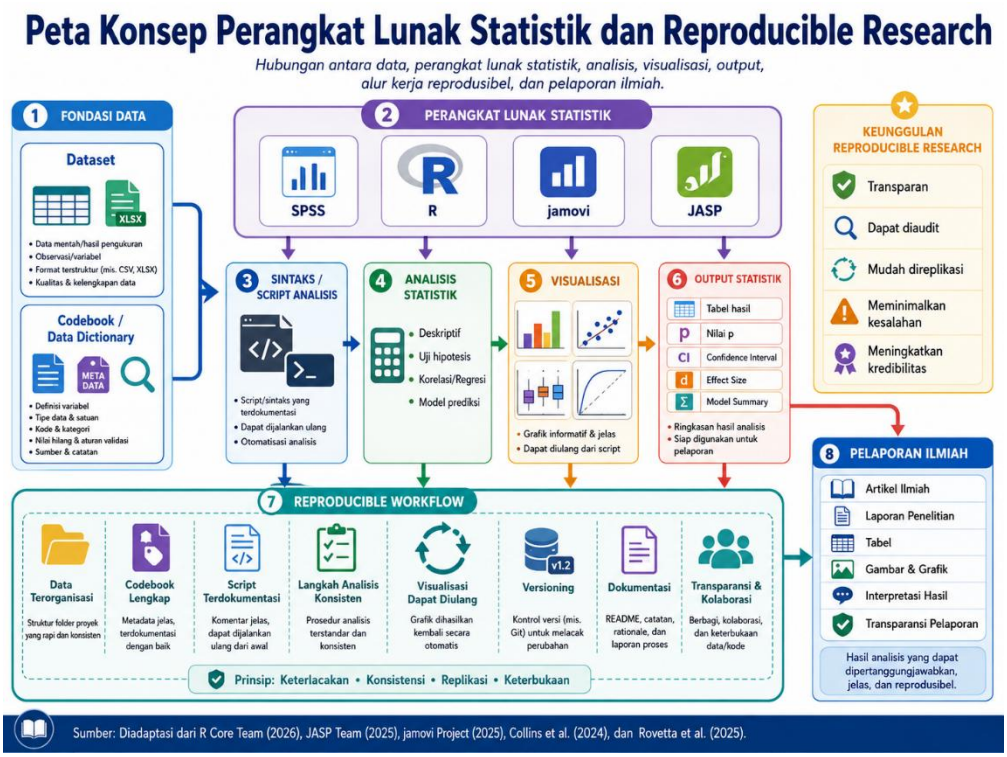
Sounderajah, V., *et al*. (2025). The STARD-AI reporting guideline for diagnostic accuracy studies evaluating AI systems. *Nature Medicine*.

BAB 9
APLIKASI PERANGKAT LUNAK STATISTIK, REPRODUCIBLE RESEARCH, MANAJEMEN DATA, DAN INTERPRETASI OUTPUT ANALISIS

9.1 Pendahuluan

Perangkat lunak statistik digunakan untuk membantu peneliti mengelola data, menjalankan analisis, membuat visualisasi, menyimpan sintaks, dan menghasilkan output yang dapat ditafsirkan secara ilmiah (R Core Team, 2026; JASP Team, 2025; jamovi Project, 2025). Penggunaan perangkat lunak tidak menggantikan pemahaman statistik karena pilihan menu, sintaks, dan output tetap harus ditentukan berdasarkan jenis variabel, desain penelitian, asumsi statistik, dan tujuan analisis (Chicco *et al*., 2025; Rovetta *et al*., 2025). Kesalahan umum dalam analisis berbasis perangkat lunak adalah menjalankan uji statistik tanpa memeriksa struktur data, missing data, outlier, distribusi, asumsi, dan kesesuaian desain penelitian (Ghasemi & Zahediasl, 2012; Rovetta *et al*., 2025). Oleh karena itu, kompetensi statistik modern tidak hanya mencakup kemampuan memilih uji, tetapi juga kemampuan mengelola dataset, mendokumentasikan proses analisis, dan melaporkan hasil secara transparan (Collins *et al*., 2024; Hopewell *et al*., 2025).

Bab ini membahas penggunaan perangkat lunak statistik seperti SPSS, R, jamovi, dan JASP dalam konteks penelitian kesehatan dan biomedis. R merupakan lingkungan perangkat lunak bebas untuk komputasi statistik dan grafik, sedangkan JASP dan jamovi menyediakan antarmuka grafis yang lebih mudah digunakan oleh pemula (R Core Team, 2026; JASP Team, 2025; jamovi Project, 2025). jamovi dibangun di atas bahasa statistik R dan menyediakan struktur spreadsheet dengan tipe data seperti integer, decimal, text, serta tipe pengukuran seperti nominal, ordinal, continuous, dan ID (jamovi Project, 2025). JASP menyediakan antarmuka grafis untuk analisis statistik klasik dan Bayesian, sehingga sering digunakan dalam pendidikan statistik dan penelitian terapan (JASP Team, 2025).



Gambar 9. 1. Peta konsep perangkat lunak statistik dan reproducible research

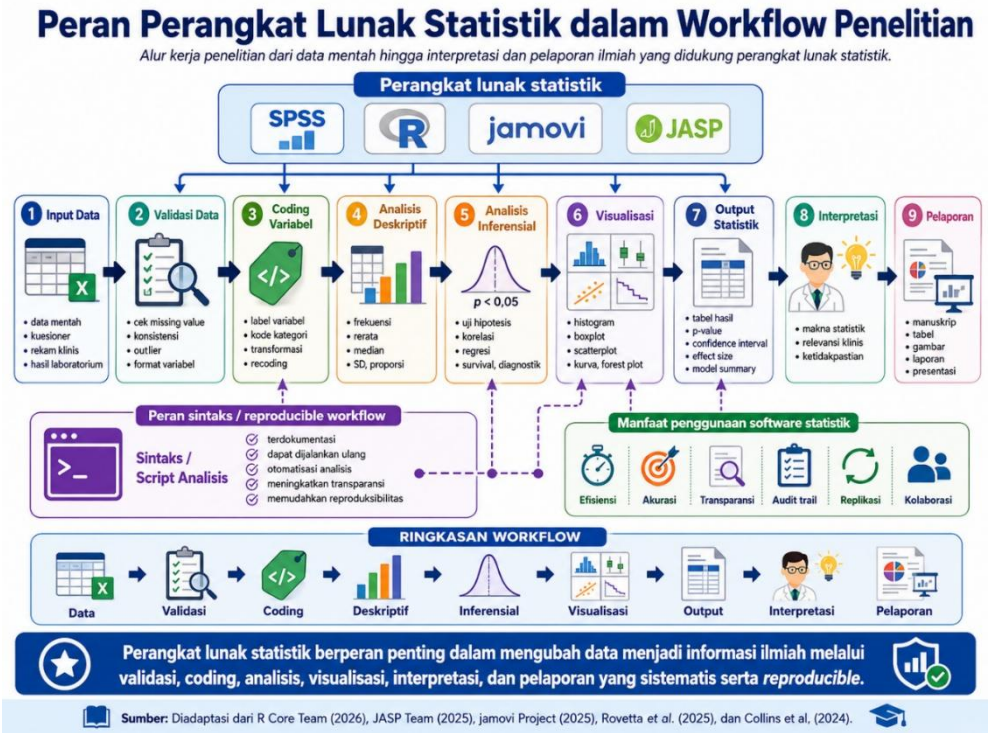
Ilustrasi menunjukkan hubungan antara dataset, codebook, SPSS, R, jamovi, JASP, sintaks analisis, visualisasi, output statistik, reproducible workflow, dan pelaporan ilmiah. **Sumber:** Diadaptasi dari R Core Team (2026), JASP Team (2025), jamovi Project (2025), Collins *et al*. (2024), dan Rovetta *et al*. (2025).

9.2 Peran Perangkat Lunak Statistik dalam Penelitian

Perangkat lunak statistik berperan dalam proses input data, pembersihan data, transformasi variabel, analisis deskriptif, analisis inferensial, visualisasi data, dan penyimpanan output penelitian (R Core Team, 2026; jamovi Project, 2025). Perangkat lunak seperti R memungkinkan analisis berbasis sintaks yang dapat direproduksi, sedangkan perangkat lunak berbasis antarmuka grafis seperti JASP dan jamovi memudahkan pengguna menjalankan analisis tanpa menulis kode panjang (R Core Team, 2026; JASP Team, 2025; jamovi Project, 2025). Dalam konteks pendidikan statistik, antarmuka grafis dapat membantu mahasiswa memahami hubungan antara variabel, pilihan uji, dan output analisis (JASP Team, 2025; jamovi Project, 2025). Namun, kemudahan

antarmuka tidak boleh membuat pengguna mengabaikan asumsi analisis dan interpretasi ilmiah (Rovetta *et al*., 2025).

Perangkat lunak statistik juga membantu dokumentasi analisis melalui penyimpanan sintaks, file proyek, log output, dan script. Dokumentasi tersebut penting karena penelitian yang baik harus memungkinkan pembaca atau peneliti lain menilai bagaimana data dibersihkan, bagaimana variabel dikodekan, bagaimana analisis dilakukan, dan bagaimana hasil diperoleh (Collins *et al*., 2024; Hopewell *et al*., 2025). Pedoman TRIPOD+AI menekankan transparansi dalam pelaporan model prediksi klinis, termasuk sumber data, prediktor, outcome, metode pemodelan, validasi, dan performa model (Collins *et al*., 2024). Dengan demikian, perangkat lunak statistik harus digunakan sebagai bagian dari workflow ilmiah yang terdokumentasi, bukan hanya alat untuk menghasilkan nilai p (Rovetta *et al*., 2025).



Gambar 9. 2. Peran perangkat lunak statistik dalam workflow penelitian

Diagram menunjukkan alur dari input data, validasi data, coding variabel, analisis deskriptif, analisis inferensial, visualisasi, output, sintaks, interpretasi, dan pelaporan. **Sumber:** Diadaptasi dari R Core Team (2026), JASP Team (2025), jamovi Project (2025), Rovetta *et al*. (2025), dan Collins *et al*. (2024).

9.3 Struktur Dataset dan Codebook

Dataset yang baik biasanya memiliki struktur satu baris untuk satu unit observasi dan satu kolom untuk satu variabel (Cooksey, 2020; Kaliyadan & Kulkarni, 2019). Struktur ini membantu peneliti menghindari kekacauan data, seperti beberapa variabel tercampur dalam satu kolom, satu subjek tersebar di banyak baris tanpa identitas yang jelas, atau kategori yang ditulis tidak konsisten (Cooksey, 2020). Codebook adalah dokumen yang menjelaskan nama variabel, definisi operasional, jenis data, kode kategori, satuan pengukuran, nilai hilang, dan aturan transformasi variabel (Cooksey, 2020; Rovetta *et al*., 2025). Codebook penting karena analisis statistik tidak dapat direproduksi dengan baik jika definisi variabel dan aturan pengkodean tidak terdokumentasi (Collins *et al*., 2024).

Dalam perangkat lunak seperti jamovi, variabel dapat didefinisikan berdasarkan tipe data dan tipe pengukuran, sehingga pengguna perlu menandai apakah variabel bersifat nominal, ordinal, continuous, atau ID (jamovi Project, 2025). Penentuan tipe pengukuran yang benar penting karena perangkat lunak dapat menampilkan pilihan analisis dan ringkasan statistik yang berbeda berdasarkan tipe variabel (jamovi Project, 2025). Misalnya, variabel nominal lebih tepat diringkas dengan frekuensi dan persentase, sedangkan variabel continuous dapat diringkas dengan mean-simpangan baku atau median-interquartile range sesuai distribusinya (Kaliyadan & Kulkarni, 2019; Ghasemi & Zahediasl, 2012). Dengan demikian, tahap definisi variabel dalam software merupakan keputusan metodologis, bukan sekadar pengaturan teknis (Kaliyadan & Kulkarni, 2019).

Tabel 9. 1. Elemen codebook penelitian

Elemen codebook	Fungsi	Contoh
Nama variabel	Identitas kolom data	age, sex, sbp, group
Definisi operasional	Menjelaskan apa yang diukur	Tekanan darah sistolik dalam mmHg
Tipe data	Menentukan bentuk data	Numerik, teks, tanggal
Skala pengukuran	Menentukan analisis	Nominal, ordinal, kontinu
Kode kategori	Menjelaskan arti kode	0 = kontrol, 1 = intervensi

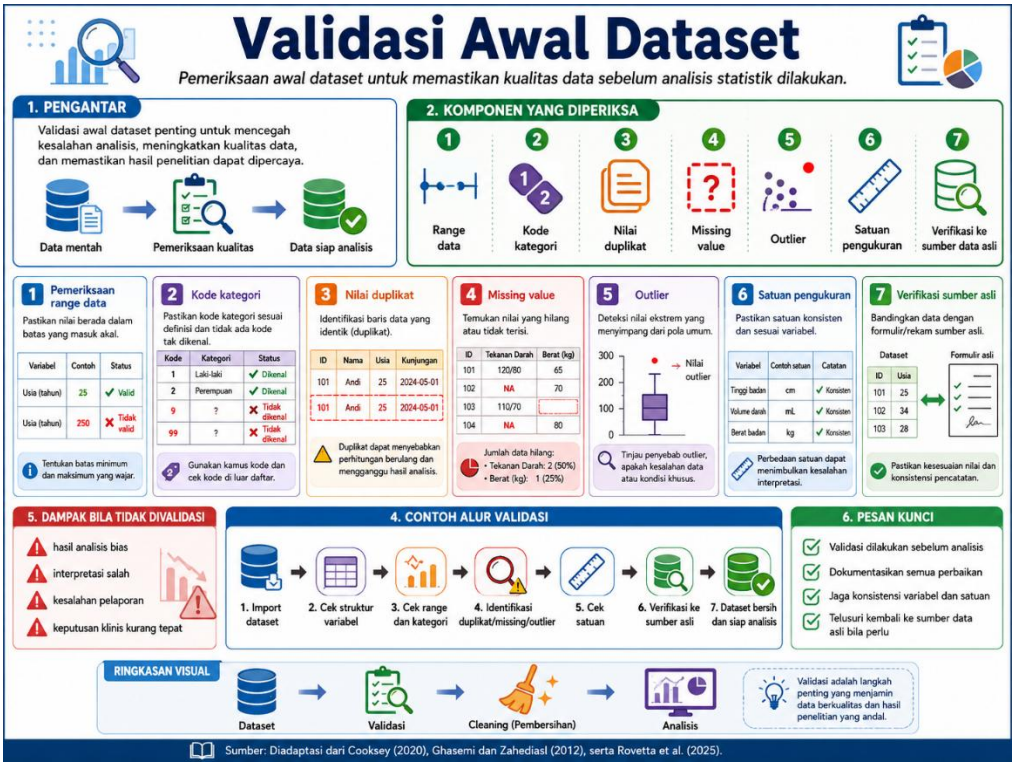
Elemen codebook	Fungsi	Contoh
Nilai hilang	Menjelaskan missing value	99 = tidak menjawab
Satuan	Menjaga interpretasi	mg/dL, mmHg, kg/m^2

Sumber tabel: Diadaptasi dari Cooksey (2020), Kaliyadan dan Kulkarni (2019), jamovi Project (2025), dan Collins *et al*. (2024).

9.4 Input Data dan Validasi Awal

Input data adalah tahap memasukkan hasil pengukuran ke dalam format digital yang dapat dianalisis oleh perangkat lunak statistik (Cooksey, 2020). Tahap ini harus dilakukan secara hati-hati karena kesalahan input dapat menghasilkan nilai ekstrem palsu, kategori tidak konsisten, data hilang, atau hasil analisis yang salah (Rovetta *et al*., 2025). Validasi awal dapat dilakukan dengan memeriksa range nilai, kategori yang tidak dikenal, nilai duplikat, data tidak lengkap, dan ketidaksesuaian satuan pengukuran (Cooksey, 2020). Pemeriksaan awal juga perlu memastikan bahwa setiap observasi memiliki identitas unik agar data tidak tercampur antarresponden atau antarwaktu pengukuran (Cooksey, 2020).

Validasi data harus dilakukan sebelum analisis deskriptif dan inferensial (Ghasemi & Zahediasl, 2012). Histogram, boxplot, tabel frekuensi, dan ringkasan numerik dapat membantu mendeteksi nilai yang tidak masuk akal (Cooksey, 2020). Misalnya, usia 250 tahun, tekanan darah sistolik 900 mmHg, atau kode jenis kelamin “7” pada variabel biner menunjukkan kemungkinan kesalahan input atau coding (Cooksey, 2020). Data yang mencurigakan harus ditelusuri kembali ke sumber asli sebelum diputuskan untuk diperbaiki, dihapus, atau dipertahankan sebagai nilai valid (Rovetta *et al*., 2025).



Gambar 9. 3. Validasi awal dataset

Gambar menunjukkan pemeriksaan range data, kode kategori, nilai duplikat, missing value, outlier, satuan pengukuran, dan verifikasi terhadap sumber data asli. **Sumber:** Diadaptasi dari Cooksey (2020), Ghasemi dan Zahediasl (2012), serta Rovetta *et al*. (2025).

9.5 Cleaning Data dan Penanganan Missing Value

Cleaning data adalah proses memperbaiki, menandai, atau mendokumentasikan masalah dalam dataset sebelum analisis dilakukan (Rovetta *et al*., 2025). Masalah yang sering ditemukan meliputi nilai hilang, salah ketik, kategori ganda, satuan tidak konsisten, duplikasi data, outlier, dan kesalahan format tanggal (Cooksey, 2020). Missing value perlu dibedakan dari nilai nol karena nol dapat berarti kuantitas nyata, sedangkan missing berarti data tidak tersedia (Kaliyadan & Kulkarni, 2019). Kesalahan membedakan nol dan missing dapat mengubah rerata, proporsi, standar deviasi, dan hasil uji statistik (Cooksey, 2020).

Penanganan missing value harus dijelaskan secara transparan dalam laporan penelitian (Rovetta *et al*., 2025). Jika data hilang sedikit dan terjadi secara acak, analisis complete case dapat dipertimbangkan, tetapi peneliti tetap perlu melaporkan jumlah data yang dianalisis (Rovetta *et al*., 2025). Jika data hilang banyak atau tidak acak, penghapusan

sederhana dapat menyebabkan bias dan penurunan power (Rovetta *et al*., 2025). Dalam model prediksi klinis, pedoman TRIPOD+AI menekankan perlunya pelaporan sumber data, missing data, metode pemodelan, dan validasi secara transparan (Collins *et al*., 2024).



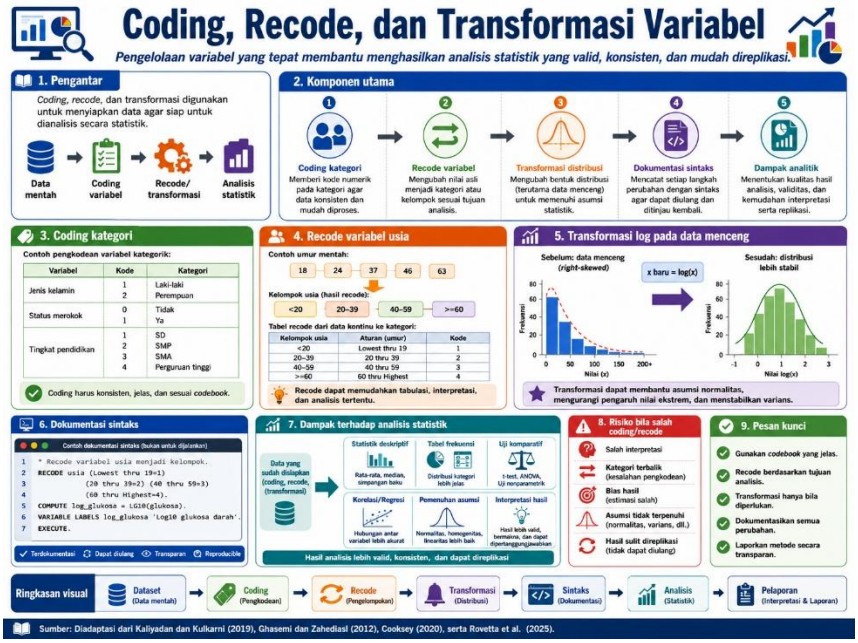
Gambar 9. 4. Cleaning data dan missing value

Diagram menunjukkan deteksi missing value, verifikasi sumber data, penandaan nilai hilang, complete case analysis, imputasi, dokumentasi keputusan, dan pelaporan transparan. **Sumber:** Diadaptasi dari Rovetta *et al*. (2025), Cooksey (2020), dan Collins *et al*. (2024).

9.6 Coding, Recode, dan Transformasi Variabel

Coding adalah proses memberikan kode numerik atau teks pada kategori variabel agar data dapat dianalisis secara konsisten (Kaliyadan & Kulkarni, 2019). Recode adalah proses mengubah kategori atau nilai variabel menjadi kategori baru sesuai kebutuhan analisis (Cooksey, 2020). Misalnya, usia dapat direcode menjadi kelompok usia, skor nyeri dapat direcode menjadi ringan-sedang-berat, atau status merokok dapat direcode menjadi perokok dan bukan perokok (Kaliyadan & Kulkarni, 2019). Recode harus didasarkan pada alasan klinis, teori, guideline, atau rencana analisis, bukan hanya agar hasil menjadi signifikan (Rovetta *et al*., 2025).

Transformasi variabel dapat dilakukan untuk mengurangi skewness, menstabilkan varians, atau memenuhi asumsi model statistik (Ghasemi & Zahediasl, 2012). Transformasi log sering digunakan pada data yang menceng ke kanan seperti biaya, lama rawat inap, atau konsentrasi biomarker tertentu (Ghasemi & Zahediasl, 2012). Namun, hasil pada skala transformasi perlu ditafsirkan dengan hati-hati karena makna klinisnya tidak selalu langsung (Rovetta *et al*., 2025). Semua proses coding, recode, dan transformasi sebaiknya dicatat dalam sintaks atau log analisis agar dapat ditelusuri ulang (Collins *et al*., 2024).



Gambar 9. 5. Coding, recode, dan transformasi variable

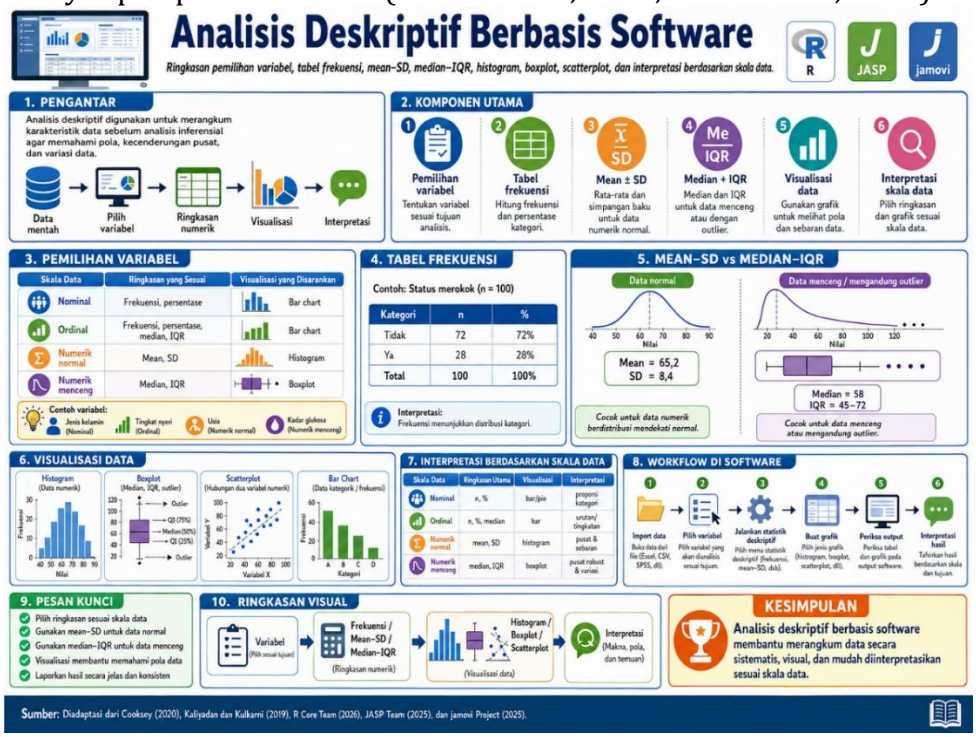
Gambar menunjukkan pengkodean kategori, recode variabel usia, transformasi log pada data menceng, dokumentasi sintaks, dan dampak terhadap analisis statistik. **Sumber:** Diadaptasi dari Kaliyadan dan Kulkarni (2019), Ghasemi dan Zahediasl (2012), Cooksey (2020), serta Rovetta *et al*. (2025).

9.7 Analisis Deskriptif dengan Perangkat Lunak Statistik

Analisis deskriptif merupakan langkah awal untuk memahami karakteristik data sebelum analisis inferensial (Cooksey, 2020). Perangkat lunak statistik dapat menghasilkan tabel frekuensi untuk variabel kategorik, ringkasan mean-simpangan baku untuk variabel numerik normal, serta median-interquartile range untuk variabel numerik yang menceng (Kaliyadan & Kulkarni, 2019; Ghasemi & Zahediasl, 2012). Visualisasi seperti histogram, boxplot, bar chart, dan

scatterplot membantu peneliti memahami distribusi data, outlier, dan pola hubungan antarvariabel (Cooksey, 2020). Output deskriptif harus ditafsirkan berdasarkan skala pengukuran dan distribusi data, bukan hanya disalin langsung dari software (Cooksey, 2020).

Dalam jamovi dan JASP, analisis deskriptif dapat dilakukan melalui antarmuka grafis sehingga pengguna dapat memilih variabel dan statistik ringkasan dengan mudah (JASP Team, 2025; jamovi Project, 2025). R dapat menjalankan analisis deskriptif melalui fungsi dasar atau paket tambahan, dan keunggulannya adalah proses analisis dapat disimpan dalam script yang mudah diulang (R Core Team, 2026). JASP juga menyediakan data library yang dapat digunakan untuk pembelajaran berbagai analisis statistik (JASP Team, 2025). Dengan demikian, perangkat lunak yang berbeda dapat menghasilkan analisis serupa, tetapi reproducibility dan dokumentasi tetap bergantung pada cara pengguna menyimpan proses analisis (R Core Team, 2026; Collins *et al*., 2024).



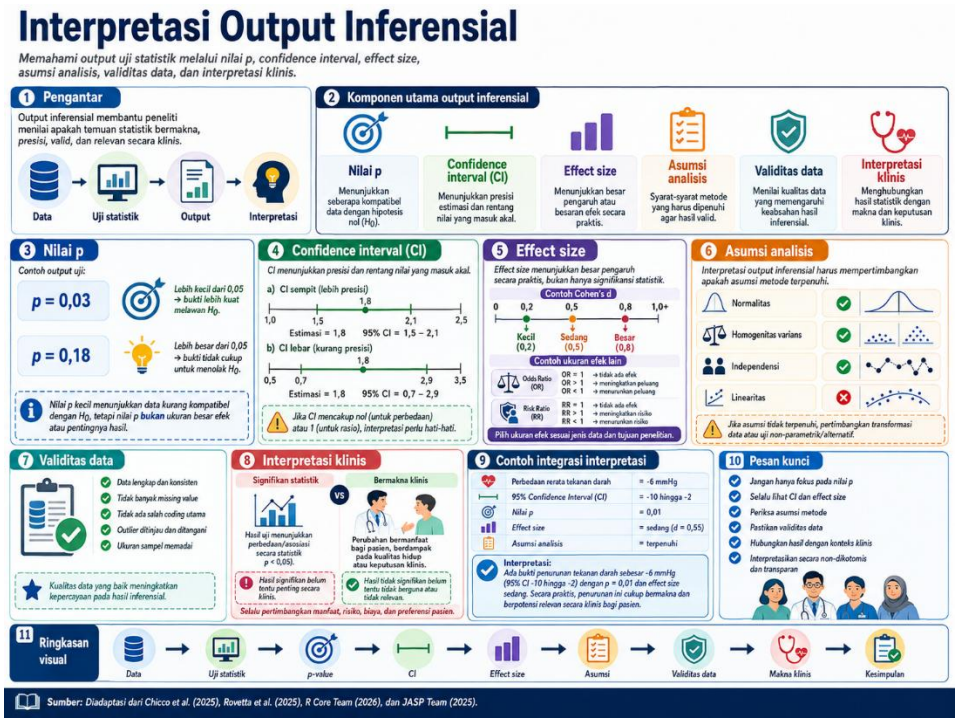
Gambar 9. 6. Analisis deskriptif berbasis software

Diagram menunjukkan pemilihan variabel, tabel frekuensi, mean-SD, median-IQR, histogram, boxplot, scatterplot, dan interpretasi berdasarkan skala data. **Sumber:** Diadaptasi dari Cooksey (2020), Kaliyadan dan Kulkarni (2019), R Core Team (2026), JASP Team (2025), dan jamovi Project (2025).

9.8 Analisis Inferensial dengan Software

Analisis inferensial dalam perangkat lunak statistik meliputi uji perbedaan kelompok, uji asosiasi, korelasi, regresi, analisis survival, dan model prediksi (Chicco *et al*., 2025; Collins *et al*., 2024). Software dapat menghitung statistik uji, nilai p, confidence interval, ukuran efek, dan grafik analisis secara cepat (R Core Team, 2026; JASP Team, 2025; jamovi Project, 2025). Namun, output software hanya benar secara teknis jika input data, jenis variabel, asumsi, dan pemilihan uji sudah benar (Rovetta *et al*., 2025). Oleh karena itu, peneliti perlu memahami apakah desain penelitian bersifat independen atau berpasangan, apakah outcome numerik atau kategorik, dan apakah asumsi parametrik terpenuhi (Chicco *et al*., 2025).

Dalam analisis inferensial, peneliti sebaiknya tidak hanya mengambil nilai p dari output software (Rovetta *et al*., 2025). Pelaporan yang baik harus mencantumkan estimasi efek, confidence interval, nilai p, ukuran efek, serta interpretasi klinis atau substantif (Rovetta *et al*., 2025). Jika output menunjukkan $p < 0,05$, peneliti tetap perlu menilai besar efek dan kemungkinan bias (Rovetta *et al*., 2025). Jika output menunjukkan $p > 0,05$, peneliti tidak boleh langsung menyimpulkan tidak ada efek tanpa melihat confidence interval dan kecukupan ukuran sampel (Rovetta *et al*., 2025).



Gambar 9. 7. Interpretasi output inferensial

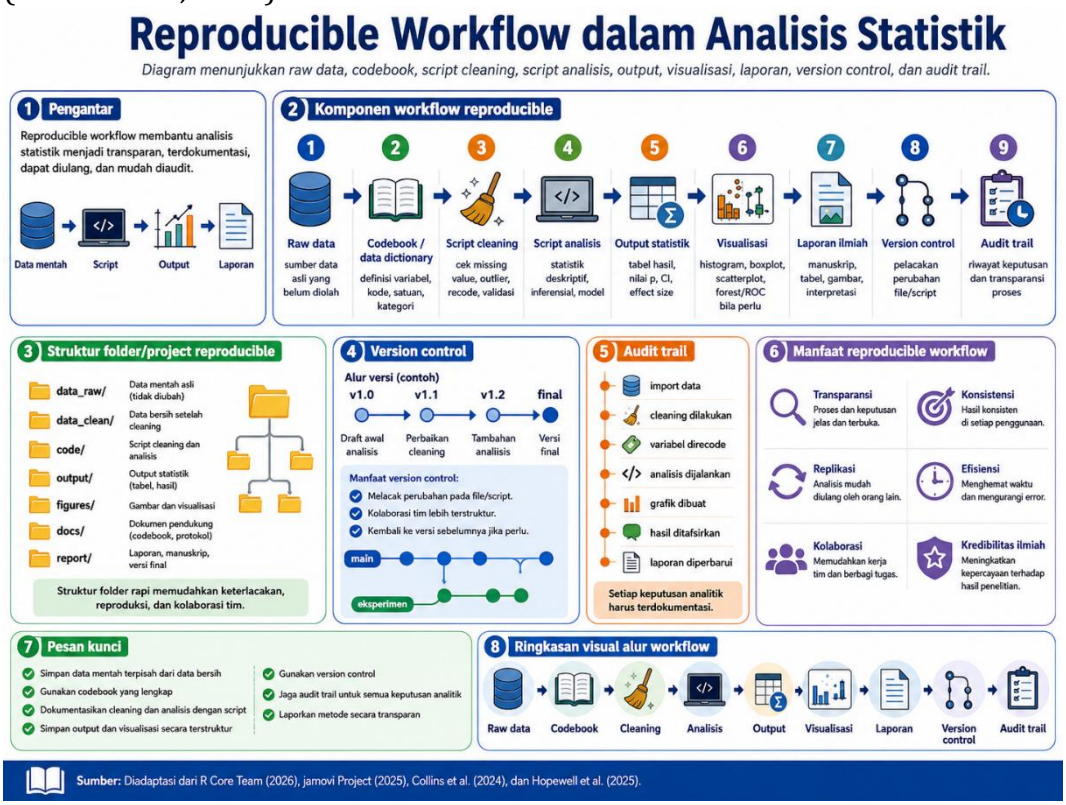
Gambar menunjukkan output uji statistik, nilai p, confidence interval, effect size, asumsi analisis, validitas data, dan interpretasi klinis. **Sumber:** Diadaptasi dari Chicco *et al*. (2025), Rovetta *et al*. (2025), R Core Team (2026), dan JASP Team (2025).

9.9 Sintaks, Script, dan Reproducible Workflow

Sintaks atau script adalah kumpulan perintah analisis yang dapat disimpan, diperiksa, dan dijalankan ulang (R Core Team, 2026). Analisis berbasis script mendukung reproducible research karena langkah-langkah analisis tidak hanya tersimpan dalam ingatan peneliti atau klik menu, tetapi terdokumentasi dalam file yang dapat diaudit (R Core Team, 2026; Collins *et al*., 2024). R sangat kuat untuk reproducible workflow karena seluruh proses dari import data, cleaning, analisis, visualisasi, hingga pelaporan dapat ditulis dalam script (R Core Team, 2026). jamovi juga dapat menyediakan kode R untuk analisis tertentu karena perangkat lunak ini dibangun di atas bahasa statistik R (jamovi Project, 2025).

Reproducible workflow penting untuk mengurangi kesalahan analisis dan meningkatkan kepercayaan terhadap hasil penelitian (Collins *et al*., 2024; Hopewell *et al*., 2025). Jika peneliti hanya mengandalkan klik menu tanpa menyimpan log, proses analisis sulit diulang dan kesalahan sulit ditelusuri (Rovetta *et al*., 2025). Reproducibility juga mendukung

kolaborasi karena anggota tim dapat memeriksa script, codebook, output, dan keputusan analisis (Collins *et al*., 2024). Dalam penelitian klinis dan model prediksi, dokumentasi lengkap diperlukan agar pembaca dapat mengevaluasi validitas, risiko bias, dan kemungkinan penerapan hasil (Collins *et al*., 2024).



Gambar 9. 8. Reproducible workflow dalam analisis statistic Diagram menunjukkan raw data, codebook, script cleaning, script analisis, output, visualisasi, laporan, version control, dan audit trail. **Sumber:** Diadaptasi dari R Core Team (2026), jamovi Project (2025), Collins *et al*. (2024), dan Hopewell *et al*. (2025).

9.10 Interpretasi Output SPSS, R, jamovi, dan JASP

Output software statistik biasanya memuat ringkasan data, statistik uji, derajat kebebasan, nilai p, confidence interval, ukuran efek, dan visualisasi (Chicco *et al*., 2025; Rovetta *et al*., 2025). Nama dan format output dapat berbeda antarsoftware, tetapi prinsip interpretasi statistik tetap sama (Chicco *et al*., 2025). Misalnya, independent t-test di SPSS, R, jamovi, atau JASP tetap harus ditafsirkan berdasarkan perbedaan rerata, confidence interval, nilai p, asumsi normalitas, dan homogenitas varians

(Chicco *et al*., 2025; Ghasemi & Zahediasl, 2012). Korelasi Pearson di software apa pun tetap harus diperiksa dengan scatterplot, outlier, linearitas, dan konteks penelitian (Wisniewski, 2024).

Perbedaan penting antarsoftware terletak pada cara pengguna menjalankan analisis dan menyimpan prosesnya. R berbasis script sehingga unggul dalam dokumentasi analisis yang dapat diulang (R Core Team, 2026). JASP dan jamovi menyediakan antarmuka grafis yang lebih mudah digunakan, tetapi pengguna tetap perlu menyimpan file proyek, output, dan keputusan analisis agar proses dapat ditelusuri (JASP Team, 2025; jamovi Project, 2025). Software tidak menentukan kualitas interpretasi karena interpretasi tetap bergantung pada pemahaman statistik, desain penelitian, dan kualitas data (Rovetta *et al*., 2025).

Tabel 9. 2. Perbandingan umum perangkat lunak statistik

Aspek	SPSS	R	jamovi	JASP
Antarmuka	Menu dan dialog	Script/kode	Spreadsheet GUI	GUI
Reproducibility	Sintaks tersedia	Sangat kuat berbasis script	Dapat menyediakan kode R	File proyek dan output
Kelebihan	Populer di pendidikan dan riset terapan	Fleksibel, gratis, kuat untuk grafik dan modelling	Mudah untuk pemula, berbasis R	Mudah untuk analisis klasik dan Bayesian
Keterbatasan	Lisensi berbayar	Kurva belajar lebih tinggi	Fitur bergantung modul	Fitur bergantung modul
Cocok untuk	Analisis terapan berbasis menu	Riset reproducible dan analisis kompleks	Pengajaran dan riset dasar-menengah	Pengajaran dan riset dasar-menengah

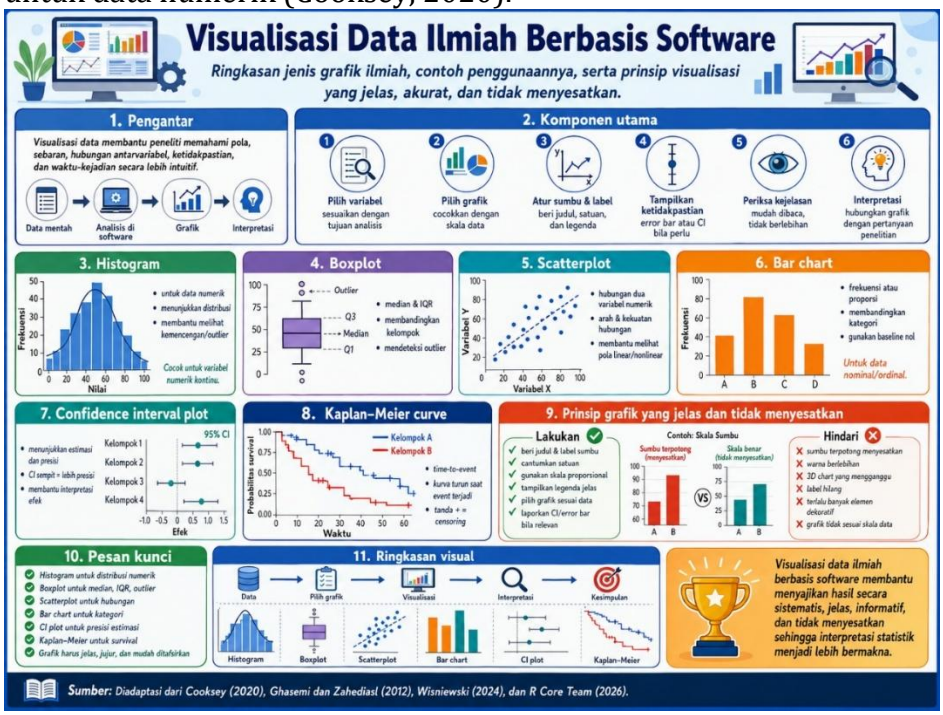
Sumber tabel: Diadaptasi dari R Core Team (2026), JASP Team (2025), jamovi Project (2025), dan prinsip reproducible research.

9.11 Visualisasi Data dan Kualitas Grafik Ilmiah

Visualisasi data membantu peneliti memahami distribusi, outlier, hubungan antarvariabel, dan pola hasil analisis (Cooksey, 2020). Grafik ilmiah yang baik harus jelas, proporsional, memiliki label sumbu, satuan,

legenda, ukuran sampel, dan caption yang menjelaskan pesan utama grafik (Cooksey, 2020). Histogram dan Q–Q plot digunakan untuk menilai distribusi numerik dan normalitas, sedangkan boxplot digunakan untuk menilai median, kuartil, dan outlier (Ghasemi & Zahediasl, 2012). Scatterplot digunakan untuk menilai hubungan dua variabel numerik sebelum korelasi atau regresi (Wisniewski, 2024).

R memiliki kemampuan grafik yang kuat karena dirancang sebagai lingkungan komputasi statistik dan grafik (R Core Team, 2026). JASP dan jamovi juga menyediakan visualisasi yang dapat dibuat melalui antarmuka grafis untuk membantu interpretasi hasil statistik (JASP Team, 2025; jamovi Project, 2025). Namun, grafik harus dibuat untuk memperjelas data, bukan memperindah laporan tanpa informasi tambahan (Cooksey, 2020). Peneliti perlu menghindari grafik yang menyesatkan, seperti skala sumbu yang terpotong tanpa alasan, pie chart yang terlalu banyak kategori, atau bar chart tanpa ukuran ketidakpastian untuk data numerik (Cooksey, 2020).



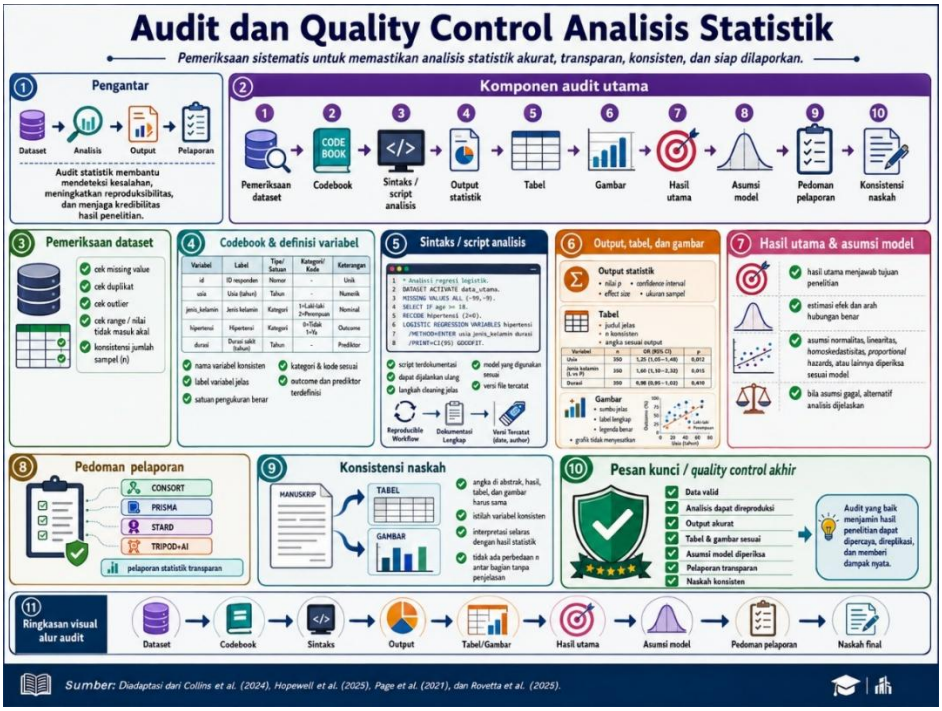
Gambar 9. 9. Visualisasi data ilmiah berbasis software

Gambar menunjukkan histogram, boxplot, scatterplot, bar chart, confidence interval plot, Kaplan–Meier curve, dan prinsip grafik yang jelas serta tidak menyesatkan. **Sumber:** Diadaptasi dari Cooksey (2020), Ghasemi dan Zahediasl (2012), Wisniewski (2024), dan R Core Team (2026).

9.12 Audit Analisis dan Quality Control

Audit analisis adalah proses memeriksa kembali dataset, codebook, sintaks, output, dan interpretasi untuk memastikan bahwa hasil penelitian konsisten dengan rencana analisis (Collins *et al*., 2024; Hopewell *et al*., 2025). Quality control perlu dilakukan sebelum hasil dimasukkan ke artikel, tesis, laporan penelitian, atau publikasi (Hopewell *et al*., 2025). Pemeriksaan ini mencakup verifikasi jumlah sampel, konsistensi kategori, missing value, outlier, pemilihan uji, asumsi model, dan kesesuaian tabel dengan output asli (Rovetta *et al*., 2025). Audit juga membantu memastikan bahwa angka dalam abstrak, tabel, gambar, dan pembahasan tidak saling bertentangan (Hopewell *et al*., 2025).

Dalam penelitian model prediksi, audit harus mencakup sumber data, definisi outcome, prediktor, metode penanganan missing data, validasi model, diskriminasi, kalibrasi, dan aksesibilitas model (Collins *et al*., 2024). Dalam randomized trial, audit pelaporan perlu mengikuti CONSORT agar alur peserta, randomisasi, intervensi, outcome, ukuran sampel, dan harm dilaporkan lengkap (Hopewell *et al*., 2025). Dalam systematic review, audit pelaporan dapat mengacu pada PRISMA agar proses pencarian, seleksi, ekstraksi, risk of bias, sintesis, dan flow diagram dilaporkan jelas (Page *et al*., 2021). Dengan demikian, quality control statistik berfungsi menjaga integritas ilmiah dari tahap data sampai publikasi (Hopewell *et al*., 2025; Page *et al*., 2021).



Gambar 9. 10. Audit dan quality control analisis statistic

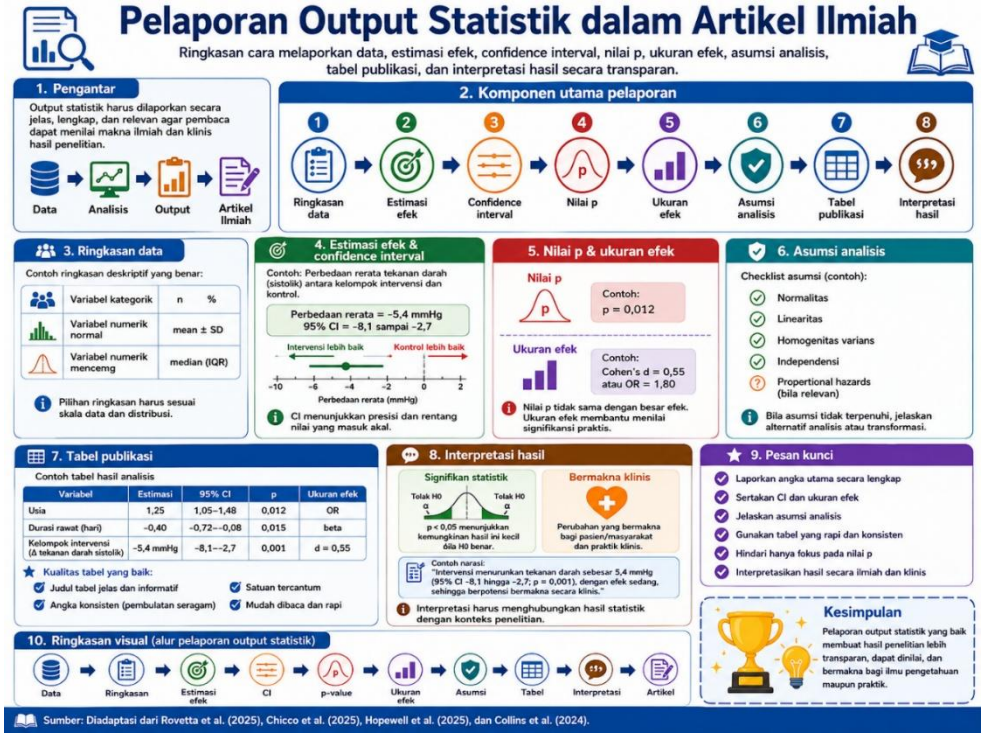
Diagram menunjukkan pemeriksaan dataset, codebook, sintaks, output, tabel, gambar, hasil utama, asumsi model, pedoman pelaporan, dan konsistensi naskah. **Sumber:** Diadaptasi dari Collins *et al*. (2024), Hopewell *et al*. (2025), Page *et al*. (2021), dan Rovetta *et al*. (2025).

9.13 Pelaporan Output Statistik dalam Artikel Ilmiah

Pelaporan output statistik harus menampilkan hasil yang relevan dengan pertanyaan penelitian dan tidak hanya menyalin seluruh output software (Rovetta *et al*., 2025). Tabel hasil sebaiknya memuat ringkasan data, estimasi efek, confidence interval, nilai p, dan ukuran efek sesuai jenis analisis (Rovetta *et al*., 2025). Untuk uji perbedaan kelompok, peneliti perlu melaporkan perbedaan rerata atau median, confidence interval, nilai p, dan effect size bila sesuai (Chicco *et al*., 2025). Untuk regresi, peneliti perlu melaporkan koefisien, odds ratio atau hazard ratio, confidence interval, nilai p, variabel yang dimasukkan, dan asumsi model (Collins *et al*., 2024).

Pelaporan yang baik harus dapat dibaca tanpa membuka output software mentah (Hopewell *et al*., 2025). Pembaca perlu memahami jumlah sampel yang dianalisis, variabel yang digunakan, metode statistik yang dipilih, hasil utama, ketidakpastian estimasi, dan interpretasi ilmiah (Rovetta *et al*., 2025). Nilai p sebaiknya tidak menjadi satu-satunya fokus

karena interval estimasi dan ukuran efek lebih membantu menilai besar serta ketidakpastian hasil (Rovetta *et al*., 2025). Jika penelitian mengikuti pedoman seperti CONSORT, PRISMA, STARD, atau TRIPOD+AI, pedoman tersebut perlu digunakan untuk memastikan kelengkapan pelaporan sesuai desain studi (Hopewell *et al*., 2025; Page *et al*., 2021; Collins *et al*., 2024).



Gambar 9. 11. Pelaporan output statistik dalam artikel ilmiah

Gambar menunjukkan ringkasan data, estimasi efek, confidence interval, nilai p, ukuran efek, asumsi analisis, tabel publikasi, dan interpretasi hasil. **Sumber:** Diadaptasi dari Rovetta *et al*. (2025), Chicco *et al*. (2025), Hopewell *et al*. (2025), dan Collins *et al*. (2024).

9.14 Kesalahan Umum dalam Penggunaan Software Statistik

Kesalahan umum pertama adalah memilih uji statistik dari menu software tanpa memahami desain penelitian (Chicco *et al*., 2025). Kesalahan kedua adalah salah mendefinisikan tipe variabel, misalnya memperlakukan data ordinal sebagai continuous tanpa alasan metodologis yang jelas (Kaliyadan & Kulkarni, 2019). Kesalahan ketiga adalah mengabaikan missing data dan outlier sebelum menjalankan uji statistik (Ghasemi & Zahediasl, 2012; Rovetta *et al*., 2025). Kesalahan keempat adalah menyalin output software secara mentah tanpa

menyaring informasi yang relevan untuk laporan ilmiah (Rovetta *et al*., 2025).

Kesalahan lain adalah tidak menyimpan sintaks atau script sehingga analisis sulit direproduksi (R Core Team, 2026; Collins *et al*., 2024). Peneliti juga sering menafsirkan nilai p secara dikotomis tanpa melihat effect size dan confidence interval (Rovetta *et al*., 2025). Dalam model prediksi, kesalahan besar adalah melaporkan performa model tanpa validasi yang memadai (Collins *et al*., 2024). Oleh karena itu, software statistik harus digunakan bersama pemahaman metodologis, dokumentasi analisis, dan pedoman pelaporan yang sesuai (Hopewell *et al*., 2025; Collins *et al*., 2024).



Gambar 9. 12. Kesalahan umum penggunaan software statistic

Diagram menunjukkan salah tipe variabel, salah uji, missing value diabaikan, outlier tidak diperiksa, output disalin mentah, nilai p dilebihkan, dan sintaks tidak disimpan. **Sumber:** Diadaptasi dari Chicco *et al*. (2025), Kaliyadan dan Kulkarni (2019), Ghasemi dan Zahediasl (2012), R Core Team (2026), dan Rovetta *et al*. (2025).

Bab ini menjelaskan bahwa perangkat lunak statistik membantu peneliti menjalankan analisis, tetapi kualitas hasil tetap bergantung pada struktur data, codebook, pemilihan uji, pemeriksaan asumsi, dokumentasi

sintaks, dan interpretasi ilmiah (R Core Team, 2026; Rovetta *et al*., 2025). R merupakan lingkungan bebas untuk komputasi statistik dan grafik, sedangkan JASP dan jamovi menyediakan antarmuka grafis yang memudahkan pengguna menjalankan analisis statistik (R Core Team, 2026; JASP Team, 2025; jamovi Project, 2025). jamovi dibangun di atas R dan menyediakan tipe data serta tipe pengukuran yang membantu pengguna mendefinisikan variabel dengan benar (jamovi Project, 2025). JASP menyediakan analisis frequentist dan Bayesian melalui antarmuka yang dirancang untuk kemudahan penggunaan (JASP Team, 2025).

Bab ini juga menekankan pentingnya reproducible research, yaitu kemampuan untuk menelusuri dan mengulang proses analisis dari data mentah sampai laporan akhir (Collins *et al*., 2024). Reproducibility dapat diperkuat melalui codebook, script, sintaks, file proyek, dokumentasi cleaning data, audit output, dan pedoman pelaporan (Collins *et al*., 2024; Hopewell *et al*., 2025). Pelaporan statistik yang baik tidak hanya menyajikan nilai p, tetapi juga estimasi efek, confidence interval, ukuran efek, asumsi model, dan interpretasi klinis (Rovetta *et al*., 2025). Dengan demikian, penguasaan software statistik harus dipahami sebagai bagian dari literasi metodologi, bukan sekadar keterampilan teknis menjalankan menu analisis (Rovetta *et al*., 2025).



Gambar 9. 13. Ringkasan perangkat lunak statistik dan reproducible research

Diagram merangkum dataset, codebook, cleaning data, coding, recode, transformasi, analisis deskriptif, analisis inferensial, visualisasi, sintaks, audit, pelaporan, dan interpretasi. **Sumber:** Diadaptasi dari R Core Team (2026), JASP Team (2025), jamovi Project (2025), Collins *et al*. (2024), Hopewell *et al*. (2025), dan Rovetta *et al*. (2025).

Review Questions Bab 9

1. **Mengapa perangkat lunak statistik tidak dapat menggantikan pemahaman statistik?**
   Jawaban: Karena software hanya menjalankan perintah, sedangkan pemilihan uji, pemeriksaan asumsi, struktur data, dan interpretasi hasil tetap membutuhkan pemahaman metodologis.
2. **Apa fungsi codebook dalam penelitian kuantitatif?**
   Jawaban: Codebook menjelaskan nama variabel, definisi operasional, tipe data, kode kategori, satuan, missing value, dan aturan transformasi data.
3. **Mengapa validasi data perlu dilakukan sebelum analisis?**
   Jawaban: Validasi data diperlukan untuk menemukan kesalahan input, kategori tidak konsisten, nilai hilang, outlier, duplikasi, dan kesalahan satuan.
4. **Apa perbedaan coding dan recode?**
   Jawaban: Coding memberikan kode pada kategori variabel, sedangkan recode mengubah nilai atau kategori variabel menjadi bentuk baru sesuai kebutuhan analisis.
5. **Apa keunggulan analisis berbasis script seperti R?**
   Jawaban: Analisis berbasis script mudah diulang, diaudit, dibagikan, dan digunakan untuk reproducible research.
6. **Apa peran jamovi dalam analisis statistik?**
   Jawaban: jamovi menyediakan spreadsheet statistik berbasis R dengan antarmuka grafis yang memudahkan analisis dan pengaturan tipe variabel.
7. **Apa peran JASP dalam pembelajaran statistik?**
   Jawaban: JASP menyediakan antarmuka grafis yang mudah digunakan untuk analisis statistik classical dan Bayesian.
8. **Mengapa output software tidak boleh disalin mentah ke artikel?**
   Jawaban: Karena artikel ilmiah harus menyajikan hasil yang relevan, terstruktur, dan ditafsirkan, bukan seluruh output teknis dari software.
9. **Apa yang dimaksud reproducible workflow?**
   Jawaban: Reproducible workflow adalah alur kerja analisis yang terdokumentasi sehingga proses dari data mentah sampai hasil akhir dapat ditelusuri dan diulang.
10. **Apa kesalahan umum dalam penggunaan software statistik?**
    Jawaban: Kesalahan umum meliputi salah mendefinisikan variabel, memilih uji tanpa memahami desain, mengabaikan

missing data dan outlier, tidak menyimpan sintaks, serta menafsirkan hasil hanya berdasarkan nilai p.

DAFTAR PUSTAKA

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*, Article 40.

Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., Ghassemi, M., Liu, X., Reitsma, J. B., van Smeden, M., Boulesteix, A. L., Camaradou, J. C., Celi, L. A., Denaxas, S., Denniston, A. K., Glocker, B., Golub, R. M., Harvey, H., Heinze, G., *et al*. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ, 385*, e078378.

Cooksey, R. W. (2020). Descriptive statistics for summarising data. In *Illustrating statistical procedures: Finding meaning in quantitative data* (pp. 61–139). Springer.

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.

Hopewell, S., Chan, A.-W., Collins, G. S., Hróbjartsson, A., Moher, D., Schulz, K. F., *et al*. (2025). CONSORT 2025 statement: Updated guideline for reporting randomised trials. *The Lancet*.

JASP Team. (2025). *JASP: A fresh way to do statistics*. JASP.

jamovi Project. (2025). *jamovi: Open statistical software for the desktop and cloud*. jamovi.

Kaliyadan, F., & Kulkarni, V. (2019). Types of variables, descriptive statistics, and sample size. *Indian Dermatology Online Journal, 10*(1), 82–86.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., *et al*. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71.

R Core Team. (2026). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rovetta, A., Piretta, L., & Mansournia, M. A. (2025). p-values and confidence intervals as compatibility measures: Guidelines for interpreting statistical studies in clinical research. *The Lancet Regional Health – Southeast Asia*.

Wisniewski, S. J. (2024). Correlation coefficient, partial correlation, Spearman rank correlation, and regression analysis. In *StatPearls*. StatPearls Publishing.

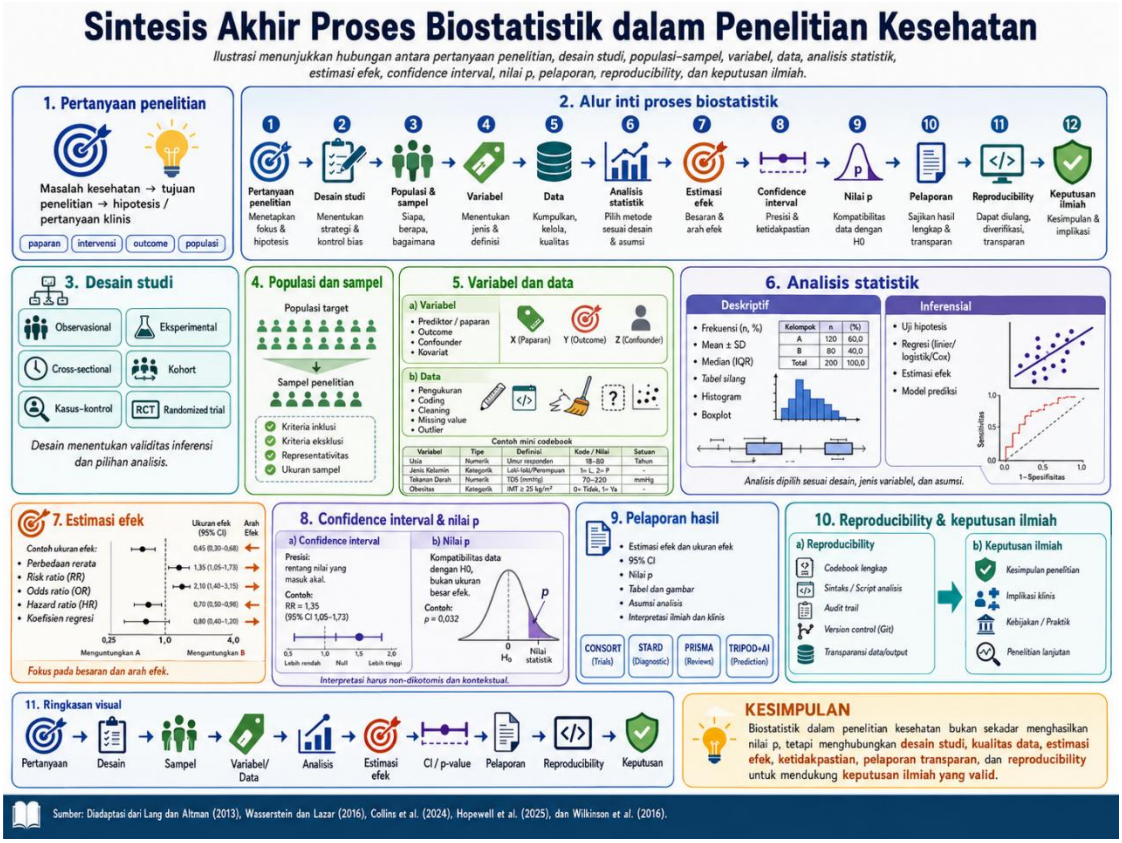
BAB 10

SINTESIS AKHIR BIOSTATISTIK: INTEGRASI DESAIN PENELITIAN, ANALISIS DATA, INTERPRETASI STATISTIK, PELAPORAN ILMIAH, DAN ETIKA REPRODUCIBLE RESEARCH

10.1 Pendahuluan

Bab 10 merupakan sintesis akhir dari pembahasan biostatistik dasar sampai lanjutan yang telah mencakup statistik deskriptif, statistik inferensial, jenis variabel, sampling, distribusi data, uji perbedaan kelompok, korelasi, regresi, analisis data kategorik, analisis diagnostik, analisis survival, model prediksi klinis, perangkat lunak statistik, dan reproducible research. Tujuan utama biostatistik bukan hanya menghasilkan nilai p, tetapi membantu peneliti merumuskan pertanyaan ilmiah, memilih desain yang tepat, mengelola data secara valid, memilih analisis yang sesuai, menafsirkan ketidakpastian, dan melaporkan hasil secara transparan. Pedoman SAMPL menekankan pentingnya pelaporan metode dan hasil statistik secara benar dalam artikel biomedis agar pembaca, editor, dan reviewer dapat menilai kualitas analisis yang digunakan.

Biostatistik modern juga harus dipahami sebagai bagian dari integritas penelitian. Hasil statistik yang tampak meyakinkan dapat menjadi keliru bila variabel salah dikodekan, sampel tidak representatif, missing data diabaikan, uji statistik tidak sesuai, asumsi model tidak diperiksa, atau hasil dilaporkan hanya berdasarkan signifikansi statistik. American Statistical Association menegaskan bahwa p-value sering disalahpahami dan tidak boleh digunakan sebagai satu-satunya ukuran kebenaran ilmiah atau kepentingan hasil penelitian. Oleh karena itu, bab ini menempatkan biostatistik sebagai proses ilmiah utuh yang menghubungkan desain, data, analisis, interpretasi, pelaporan, dan etika penelitian.



Gambar 10. 1. Sintesis akhir proses biostatistik dalam penelitian Kesehatan

Ilustrasi menunjukkan hubungan antara pertanyaan penelitian, desain studi, populasi-sampel, variabel, data, analisis statistik, estimasi efek, confidence interval, nilai p, pelaporan, reproducibility, dan keputusan ilmiah. **Sumber:** Diadaptasi dari Lang dan Altman (2013), Wasserstein dan Lazar (2016), Collins *et al*. (2024), Hopewell *et al*. (2025), dan Wilkinson *et al*. (2016).

10.2 Dari Pertanyaan Penelitian ke Rencana Analisis Statistik

Analisis statistik yang baik dimulai dari pertanyaan penelitian yang jelas. Pertanyaan penelitian menentukan desain studi, populasi target, variabel utama, outcome primer, variabel kovariat, hipotesis, dan metode analisis. Dalam uji klinis, ICH E9(R1) menekankan pentingnya estimand, yaitu definisi presisi tentang efek terapi yang ingin diestimasi berdasarkan tujuan klinis, populasi, variabel, intercurrent events, dan ringkasan populasi yang relevan.

Rencana analisis statistik perlu ditentukan sebelum data dianalisis agar keputusan analitik tidak dipengaruhi oleh hasil yang diharapkan. Rencana tersebut mencakup definisi outcome, metode pengelolaan missing data, teknik analisis utama, analisis sensitivitas, batas signifikansi, ukuran efek, dan cara pelaporan hasil. ICH E9(R1) juga menekankan bahwa tujuan klinis harus diterjemahkan menjadi pertanyaan klinis yang jelas melalui estimand yang sesuai, sehingga analisis statistik menjawab pertanyaan yang benar, bukan sekadar menjalankan uji yang tersedia.



Gambar 10. 2. Alur dari pertanyaan penelitian ke rencana analisis statistic

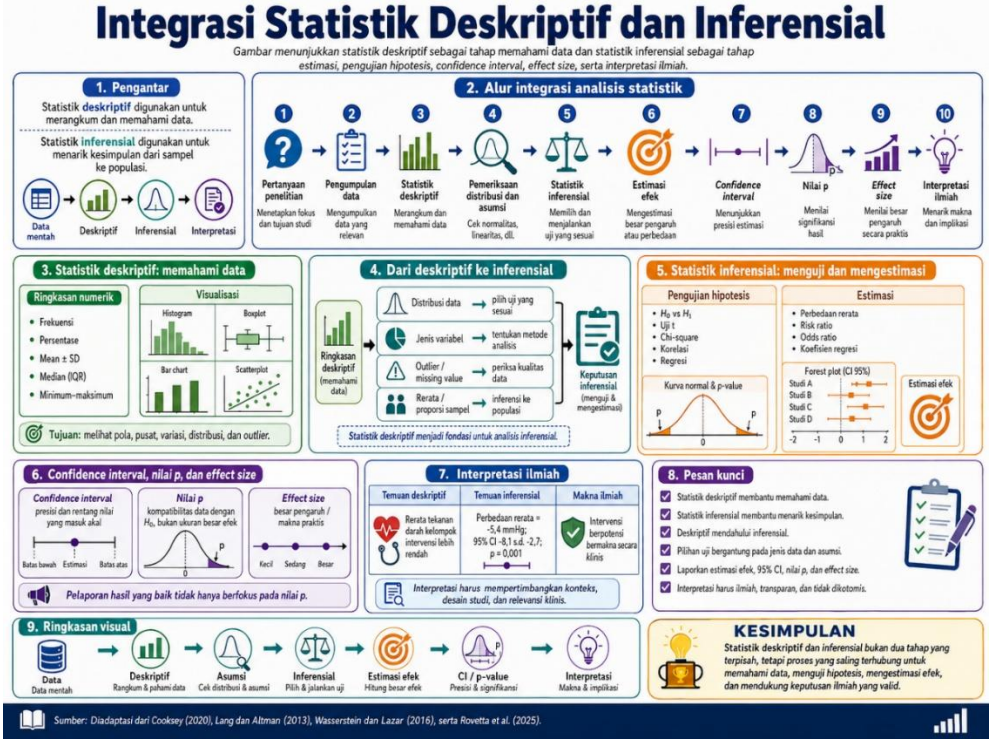
Diagram menunjukkan pertanyaan penelitian, desain studi, outcome primer, estimand, variabel prediktor, confounder, metode analisis, analisis sensitivitas, dan pelaporan hasil. **Sumber:** Diadaptasi dari ICH E9(R1) (2019/2020), Kahan *et al*. (2024), dan SAMPL Guidelines.

10.3 Integrasi Statistik Deskriptif dan Inferensial

Statistik deskriptif memberikan ringkasan awal tentang karakteristik data, sedangkan statistik inferensial membantu peneliti memperkirakan parameter populasi atau menguji hipotesis berdasarkan sampel. Statistik deskriptif diperlukan sebelum analisis inferensial karena

peneliti harus memahami jenis variabel, distribusi data, missing value, outlier, dan pola hubungan antarvariabel. Tanpa pemeriksaan deskriptif, analisis inferensial berisiko dilakukan pada data yang belum valid atau belum sesuai dengan asumsi model.

Statistik inferensial harus dilaporkan dengan estimasi efek, confidence interval, dan nilai p secara proporsional. ASA menekankan bahwa p-value tidak mengukur probabilitas bahwa hipotesis benar dan tidak mengukur besar efek atau kepentingan hasil. Dengan demikian, interpretasi statistik perlu menggabungkan besar efek, presisi estimasi, arah hubungan, plausibilitas biologis, desain penelitian, potensi bias, dan relevansi klinis.



Gambar 10. 3. Integrasi statistik deskriptif dan inferensial

Gambar menunjukkan statistik deskriptif sebagai tahap memahami data dan statistik inferensial sebagai tahap estimasi, pengujian hipotesis, confidence interval, effect size, serta interpretasi ilmiah. **Sumber:** Diadaptasi dari Cooksey (2020), Lang dan Altman (2013), Wasserstein dan Lazar (2016), serta Rovetta *et al*. (2025).

10.4 Pemilihan Uji Statistik Berdasarkan Data dan Desain

Pemilihan uji statistik harus mengikuti struktur data dan desain penelitian. Variabel kategorik dianalisis dengan frekuensi, persentase, chi-square test, Fisher's exact test, McNemar test, atau model regresi

logistik sesuai konteks. Variabel numerik dapat dianalisis dengan uji parametrik seperti t-test, ANOVA, korelasi Pearson, dan regresi linear bila asumsi terpenuhi. Data ordinal, data sangat menceng, atau data dengan outlier kuat dapat dianalisis dengan metode nonparametrik seperti Mann–Whitney, Wilcoxon, Kruskal–Wallis, Friedman, Spearman, atau Kendall.

Pemilihan uji juga harus membedakan desain independen dan berpasangan. Dua kelompok independen tidak dianalisis dengan paired t-test, sedangkan data sebelum–sesudah pada subjek yang sama tidak dianalisis dengan independent t-test. Dalam penelitian observasional, STROBE menyediakan checklist pelaporan untuk studi kohort, case-control, dan cross-sectional agar desain, peserta, variabel, bias, ukuran studi, metode statistik, dan hasil dilaporkan secara transparan.



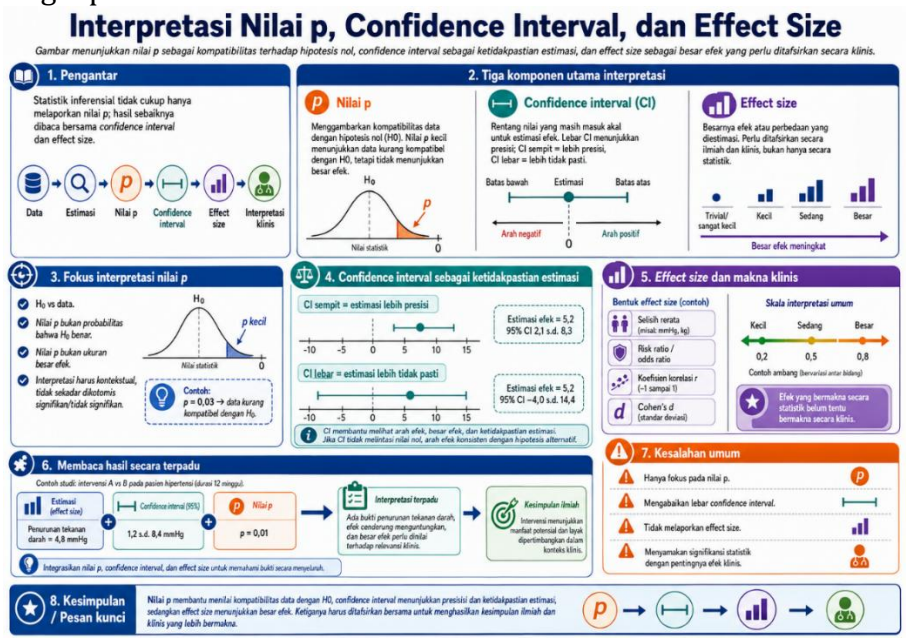
Gambar 10. 4. Matriks pemilihan uji statistik berdasarkan desain dan jenis data

Diagram menunjukkan pilihan analisis untuk data kategorik, ordinal, numerik normal, numerik tidak normal, desain independen, desain berpasangan, korelasi, regresi, diagnostik, dan survival. **Sumber:** Diadaptasi dari Chicco *et al*. (2025), STROBE Statement, SAMPL Guidelines, dan Kosin Medical Journal (2025).

10.5 Nilai p, Confidence Interval, dan Effect Size

Nilai p menunjukkan kompatibilitas data dengan hipotesis nol berdasarkan model statistik yang digunakan, tetapi tidak menunjukkan besar efek, kepentingan klinis, atau kebenaran hipotesis. ASA menyatakan bahwa p-value tidak boleh digunakan secara mekanis sebagai garis pemisah tunggal antara hasil "bermakna" dan "tidak bermakna." Oleh karena itu, laporan hasil statistik harus menjelaskan estimasi efek dan ketidakpastian, bukan hanya menampilkan $p < 0,05$.

Confidence interval memberikan informasi tentang rentang nilai parameter yang kompatibel dengan data dan model statistik. Effect size menunjukkan besarnya perbedaan, hubungan, atau asosiasi sehingga membantu pembaca menilai kepentingan praktis hasil. Pedoman SAMPL menekankan bahwa metode statistik dan hasil perlu dilaporkan cukup rinci agar pembaca dapat memahami apa yang dianalisis dan bagaimana kesimpulan dibuat. Dengan demikian, nilai p, confidence interval, dan effect size harus dipandang sebagai tiga komponen yang saling melengkapi.



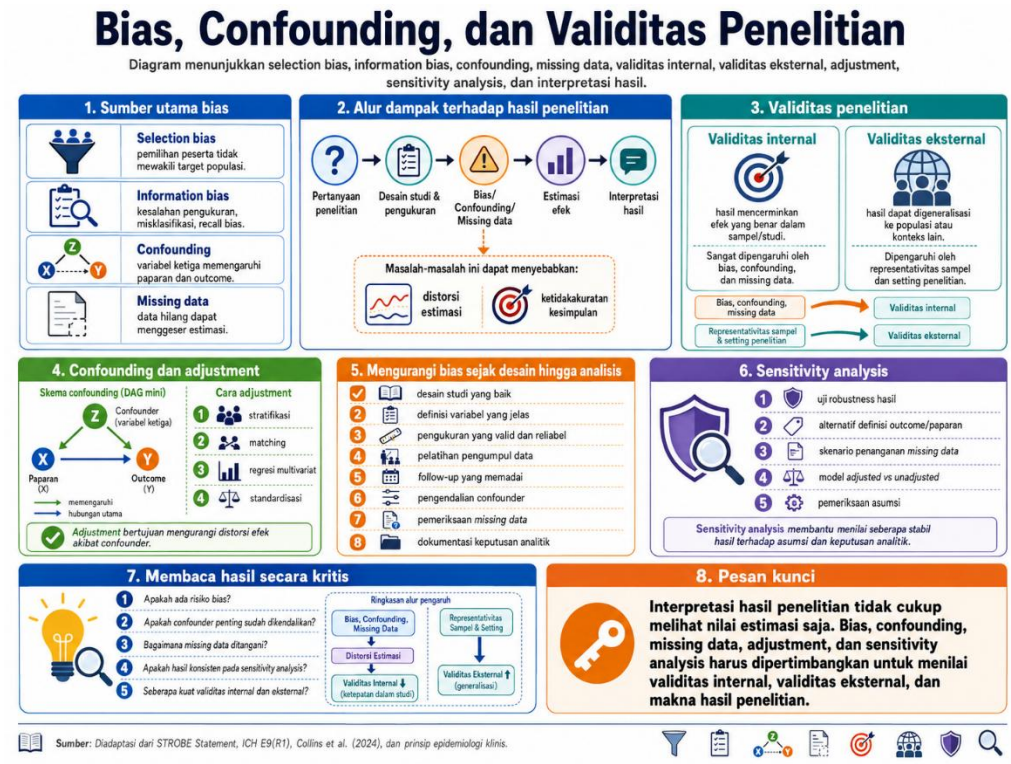
Gambar 10. 5. Interpretasi nilai p, confidence interval, dan effect size

Gambar menunjukkan nilai p sebagai kompatibilitas terhadap hipotesis nol, confidence interval sebagai ketidakpastian estimasi, dan effect size sebagai besar efek yang perlu ditafsirkan secara klinis. **Sumber:** Diadaptasi dari Wasserstein dan Lazar (2016), Lang dan Altman (2013), Rovetta *et al*. (2025), serta ASA Statement.

10.6 Bias, Confounding, dan Validitas Penelitian

Bias adalah penyimpangan sistematis yang dapat membuat hasil penelitian berbeda dari kondisi sebenarnya. Bias dapat muncul dari pemilihan sampel, pengukuran variabel, kehilangan follow-up, missing data, selective reporting, atau kesalahan analisis. Confounding terjadi ketika hubungan antara paparan dan outcome dipengaruhi oleh variabel ketiga yang berkaitan dengan keduanya. Analisis statistik dapat membantu mengontrol confounding yang diukur, tetapi tidak dapat memperbaiki semua kelemahan desain penelitian.

Validitas internal berkaitan dengan sejauh mana hasil penelitian benar untuk sampel yang diteliti, sedangkan validitas eksternal berkaitan dengan sejauh mana hasil dapat digeneralisasi ke populasi lain. STROBE menekankan bahwa laporan studi observasional harus mencakup informasi tentang desain, setting, peserta, variabel, sumber data, bias, ukuran studi, metode statistik, dan hasil agar pembaca dapat menilai validitas penelitian. Dengan demikian, interpretasi statistik harus selalu mempertimbangkan potensi bias dan confounding, bukan hanya output software.



Gambar 10. 6. Bias, confounding, dan validitas penelitian

Diagram menunjukkan selection bias, information bias, confounding, missing data, validitas internal, validitas eksternal, adjustment, sensitivity analysis, dan interpretasi hasil. **Sumber:** Diadaptasi dari STROBE Statement, ICH E9(R1), Collins *et al*. (2024), dan prinsip epidemiologi klinis.

10.7 Reproducible Research dan FAIR Data

Reproducible research menuntut agar proses penelitian dapat ditelusuri dari data mentah, codebook, cleaning data, sintaks analisis, output, visualisasi, sampai laporan akhir. Reproducibility membantu mencegah kesalahan analisis, meningkatkan kepercayaan hasil, dan memungkinkan peneliti lain memeriksa atau mengulang proses yang sama. Dalam konteks data ilmiah, FAIR Principles menekankan bahwa data dan aset digital sebaiknya Findable, Accessible, Interoperable, dan Reusable.

FAIR tidak selalu berarti semua data harus dibuka tanpa batas, terutama dalam penelitian kesehatan yang memuat informasi pribadi atau sensitif. FAIR lebih menekankan pengelolaan data yang memungkinkan penemuan, akses terkontrol, interoperabilitas, dan

penggunaan ulang sesuai etika serta regulasi. Wilkinson *et al*. menjelaskan bahwa FAIR memberi perhatian pada kemampuan mesin untuk menemukan dan menggunakan data, sekaligus mendukung penggunaan ulang oleh manusia. Oleh karena itu, penelitian biomedis modern perlu menggabungkan reproducibility, perlindungan privasi, metadata yang jelas, dan tata kelola data yang bertanggung jawab.



Gambar 10. 7. Reproducible workflow dan FAIR data dalam penelitian Kesehatan

Gambar menunjukkan raw data, metadata, codebook, script, version control, audit trail, FAIR principles, data sharing terkontrol, privasi, dan reproduksi analisis. **Sumber:** Diadaptasi dari Wilkinson *et al*. (2016), Collins *et al*. (2024), Hopewell *et al*. (2025), dan R Core Team principles.

10.8 Pedoman Pelaporan Berdasarkan Desain Studi

Setiap desain penelitian memiliki pedoman pelaporan yang sesuai. Randomized trials sebaiknya mengikuti CONSORT, studi observasional mengikuti STROBE, studi akurasi diagnostik mengikuti STARD, systematic review dan meta-analysis mengikuti PRISMA, dan model prediksi klinis mengikuti TRIPOD+AI. CONSORT 2025 terdiri atas checklist 30 item dan diagram alur peserta yang membantu pelaporan uji acak secara lengkap.

PRISMA 2020 menggantikan PRISMA 2009 dan menyediakan checklist 27 item serta diagram alur untuk systematic review dan meta-analysis. TRIPOD+AI menyediakan checklist 27 item untuk pelaporan studi model prediksi klinis yang menggunakan regresi atau machine learning, baik untuk diagnosis, prognosis, monitoring, maupun skrining. Dengan demikian, pedoman pelaporan bukan aksesori administratif, tetapi alat untuk meningkatkan transparansi, replikasi, dan kualitas interpretasi penelitian.

Tabel 10. 1. Pedoman pelaporan utama dalam penelitian kesehatan

Desain penelitian	Pedoman pelaporan	Fokus utama
Randomized trial	CONSORT 2025	Randomisasi, alokasi, intervensi, outcome, flow peserta, harm
Observasional	STROBE	Cohort, case-control, cross-sectional, bias, variabel, metode statistik
Akurasi diagnostik	STARD	Index test, reference standard, tabel 2×2, sensitivitas, spesifisitas
Systematic review/meta-analysis	PRISMA 2020	Pencarian, seleksi studi, risk of bias, sintesis, flow diagram
Model prediksi klinis	TRIPOD+AI	Prediktor, outcome, pemodelan, validasi, diskriminasi, kalibrasi
Statistik dasar biomedis	SAMPL	Pelaporan metode statistik dan hasil dasar

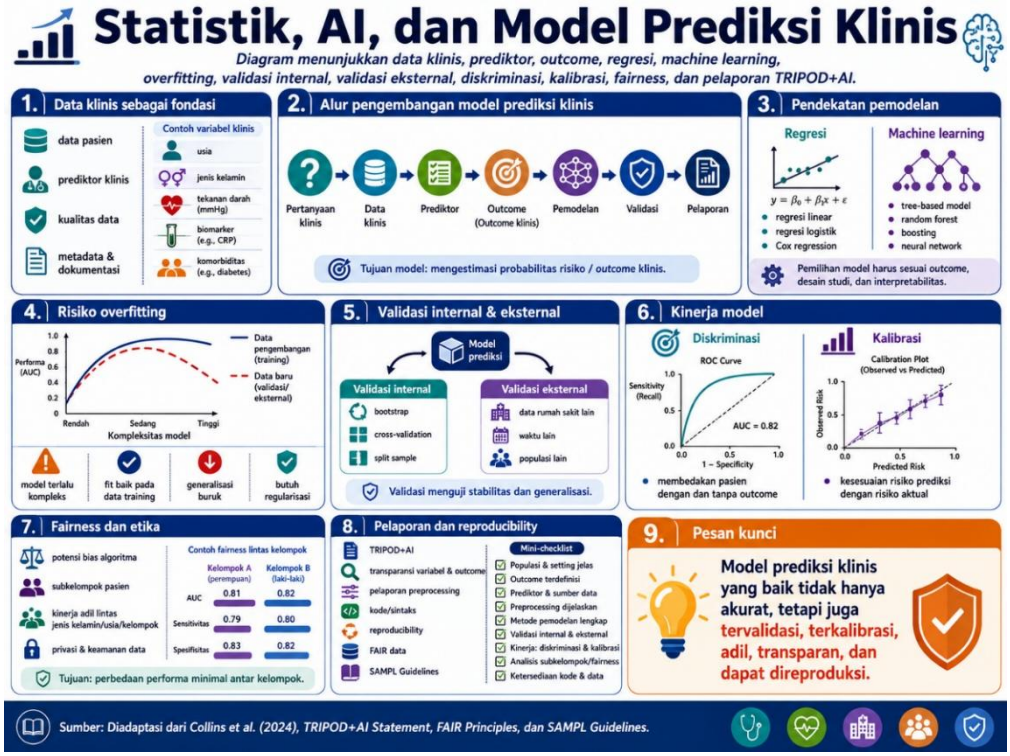
Sumber tabel: Diadaptasi dari CONSORT 2025, STROBE, STARD, PRISMA 2020, TRIPOD+AI, dan SAMPL Guidelines.

10.9 Statistik dalam Era Artificial Intelligence dan Model Prediksi

Perkembangan artificial intelligence dan machine learning memperluas penggunaan statistik dalam penelitian kesehatan. Model prediksi dapat dikembangkan untuk diagnosis, prognosis, skrining, monitoring, triase, atau penentuan risiko komplikasi. Namun, model yang tampak akurat pada data pengembangan dapat gagal pada populasi baru bila overfitting, bias data, kalibrasi buruk, atau validasi eksternal tidak dilakukan.

TRIPOD+AI memberikan panduan pelaporan harmonis untuk model prediksi klinis, baik yang menggunakan regresi maupun machine

learning. Pedoman ini menggantikan TRIPOD 2015 dan menekankan transparansi pada studi pengembangan, validasi, atau updating model prediksi. Model prediksi yang baik harus dilaporkan dengan sumber data, outcome, prediktor, missing data, metode pemodelan, validasi, diskriminasi, kalibrasi, dan potensi penggunaan klinis.



Gambar 10. 8. Statistik, AI, dan model prediksi klinis

Diagram menunjukkan data klinis, prediktor, outcome, regresi, machine learning, overfitting, validasi internal, validasi eksternal, diskriminasi, kalibrasi, fairness, dan pelaporan TRIPOD+AI. **Sumber:** Diadaptasi dari Collins *et al*. (2024), TRIPOD+AI Statement, FAIR Principles, dan SAMPL Guidelines.

10.10 Etika Statistik dan Tanggung Jawab Peneliti

Etika statistik mencakup kejujuran dalam perencanaan analisis, pengelolaan data, pelaporan hasil, dan interpretasi ketidakpastian. Peneliti tidak boleh memilih uji statistik hanya untuk memperoleh hasil signifikan, mengubah outcome setelah melihat hasil tanpa pelaporan, menghapus outlier tanpa alasan, atau hanya melaporkan analisis yang mendukung hipotesis. Praktik seperti p-hacking, selective reporting, dan HARKing dapat merusak validitas ilmiah.

Etika statistik juga menuntut transparansi terhadap keterbatasan penelitian. Jika ukuran sampel kecil, confidence interval lebar, missing data banyak, asumsi model tidak sepenuhnya terpenuhi, atau generalisasi terbatas, hal tersebut perlu disampaikan secara jelas. ASA menekankan bahwa praktik ilmiah yang baik membutuhkan teknik yang dipilih secara tepat, analisis yang dilakukan dengan benar, dan interpretasi yang merepresentasikan ketidakpastian secara layak. Dengan demikian, statistik bukan hanya keterampilan teknis, tetapi juga komitmen etik terhadap kebenaran ilmiah.



Gambar 10. 9. Etika statistik dalam penelitian biomedis

Gambar menunjukkan preregistration, analysis plan, transparansi data, pelaporan lengkap, penghindaran p-hacking, pelaporan uncertainty, dan tanggung jawab interpretasi. **Sumber:** Diadaptasi dari ASA Statement, SAMPL Guidelines, CONSORT 2025, PRISMA 2020, dan FAIR Principles.

10.11 Audit Statistik sebelum Publikasi

Audit statistik sebelum publikasi diperlukan untuk memastikan bahwa data, analisis, output, tabel, gambar, dan interpretasi konsisten. Audit dimulai dari memeriksa jumlah sampel, definisi variabel, missing data, outlier, asumsi analisis, dan kesesuaian uji statistik. Setelah itu,

peneliti perlu memeriksa apakah hasil dalam abstrak, tabel, gambar, dan pembahasan memiliki angka yang sama dan tidak saling bertentangan.

Checklist pelaporan dapat digunakan sebagai alat audit. CONSORT menggunakan checklist dan flow diagram untuk randomized trials, PRISMA menggunakan checklist serta flow diagram untuk systematic review, STROBE menyediakan checklist untuk studi observasional, dan TRIPOD+AI menyediakan checklist untuk model prediksi klinis. Dengan demikian, audit statistik bukan sekadar pemeriksaan akhir, tetapi proses sistematis untuk meningkatkan kredibilitas dan keterbacaan laporan penelitian.



Gambar 10. 10. Audit statistik sebelum publikasi

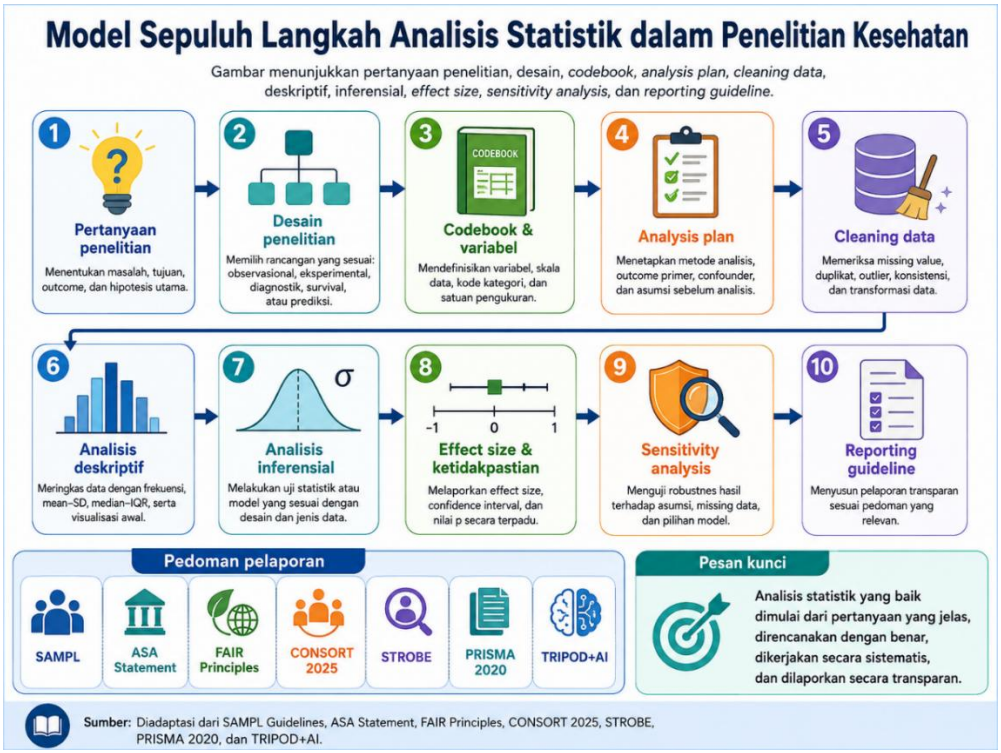
Diagram menunjukkan pemeriksaan dataset, codebook, analisis deskriptif, uji inferensial, output software, tabel, gambar, pedoman pelaporan, konsistensi angka, dan interpretasi akhir. **Sumber:** Diadaptasi dari CONSORT 2025, STROBE, PRISMA 2020, TRIPOD+AI, dan SAMPL Guidelines.

10.12 Model Praktis Akhir: Sepuluh Langkah Analisis Statistik

Model praktis akhir analisis statistik dapat dirumuskan dalam sepuluh langkah. Pertama, rumuskan pertanyaan penelitian dan outcome utama. Kedua, tentukan desain studi dan populasi target. Ketiga, definisikan variabel dan codebook. Keempat, susun rencana analisis

statistik sebelum melihat hasil. Kelima, lakukan validasi dan cleaning data. Keenam, lakukan analisis deskriptif dan visualisasi awal. Ketujuh, pilih uji statistik berdasarkan jenis data, desain, dan asumsi. Kedelapan, laporkan estimasi efek, confidence interval, nilai p, dan ukuran efek. Kesembilan, lakukan analisis sensitivitas atau validasi bila diperlukan. Kesepuluh, gunakan pedoman pelaporan sesuai desain studi.

Model ini selaras dengan prinsip SAMPL yang menekankan pelaporan metode dan hasil statistik secara jelas, prinsip ASA yang menekankan interpretasi p-value secara hati-hati, prinsip FAIR yang menekankan tata kelola data ilmiah, serta pedoman pelaporan seperti CONSORT, STROBE, PRISMA, dan TRIPOD+AI yang menekankan transparansi penelitian. Dengan demikian, analisis statistik yang baik adalah proses yang terencana, terdokumentasi, transparan, dapat diaudit, dan relevan secara ilmiah.

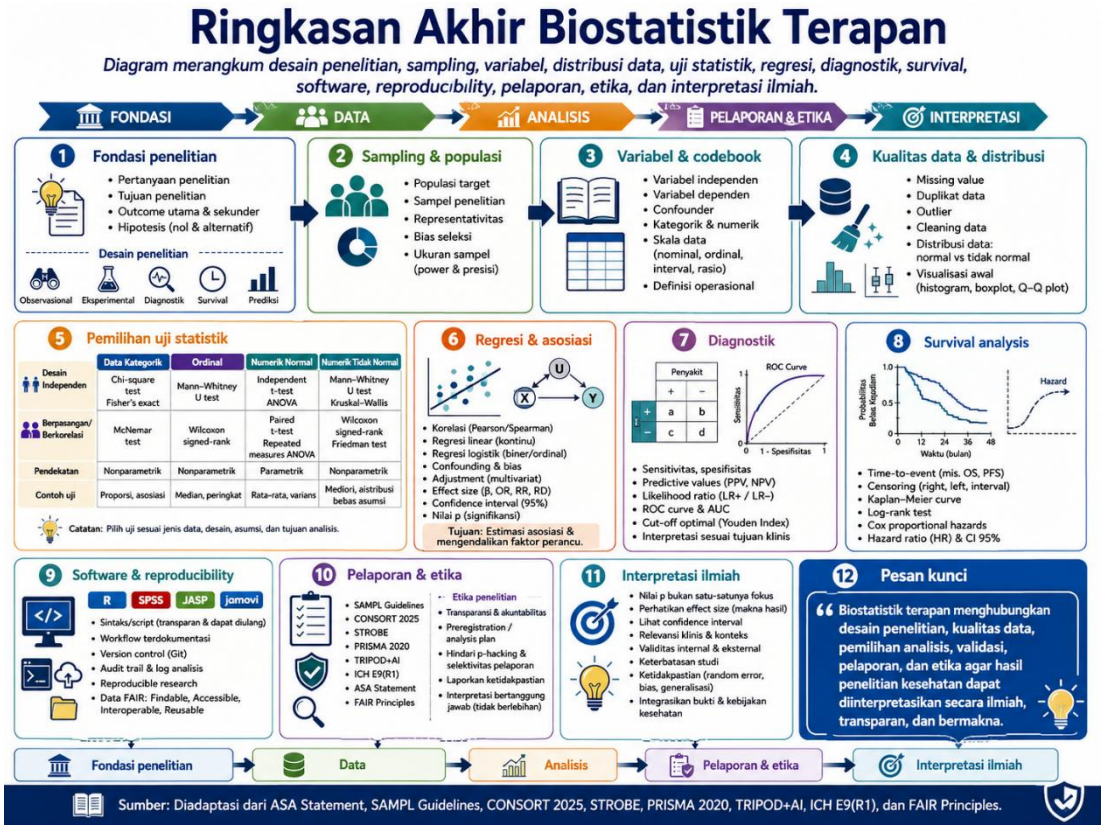


Gambar 10. 11. Model sepuluh langkah analisis statistik dalam penelitian Kesehatan

Gambar menunjukkan pertanyaan penelitian, desain, codebook, analysis plan, cleaning data, deskriptif, inferensial, effect size, sensitivity analysis, dan reporting guideline. **Sumber:** Diadaptasi dari SAMPL Guidelines, ASA Statement, FAIR Principles, CONSORT 2025, STROBE, PRISMA 2020, dan TRIPOD+AI.

Bab ini menyimpulkan bahwa biostatistik adalah kerangka penalaran ilmiah yang menghubungkan desain penelitian, struktur data, analisis statistik, interpretasi hasil, dan pelaporan ilmiah. Statistik deskriptif membantu memahami data, statistik inferensial membantu mengestimasi efek dan ketidakpastian, uji parametrik dan nonparametrik dipilih berdasarkan karakteristik data, regresi membantu pemodelan hubungan, analisis diagnostik menilai performa tes, survival analysis menangani time-to-event data, dan model prediksi mendukung estimasi risiko individual. Semua analisis tersebut harus didasarkan pada pertanyaan penelitian yang jelas dan rencana analisis yang dapat dipertanggungjawabkan.

Bab ini juga menegaskan bahwa kualitas penelitian tidak hanya bergantung pada software atau rumus statistik. Kualitas penelitian bergantung pada transparansi, reproducibility, tata kelola data, audit analisis, pelaporan lengkap, dan interpretasi yang tidak melebih-lebihkan hasil. CONSORT 2025, STROBE, STARD, PRISMA 2020, TRIPOD+AI, SAMPL, ASA Statement, ICH E9(R1), dan FAIR Principles memberikan fondasi penting untuk membangun praktik statistik yang lebih terbuka, valid, dan bertanggung jawab. Dengan demikian, biostatistik terbaik bukan hanya analisis yang menghasilkan angka, tetapi analisis yang membantu menghasilkan kesimpulan ilmiah yang jujur, dapat diuji, dan bermanfaat bagi kesehatan masyarakat.



Gambar 10. 12. Ringkasan akhir biostatistik terapan

Diagram merangkum desain penelitian, sampling, variabel, distribusi data, uji statistik, regresi, diagnostik, survival, software, reproducibility, pelaporan, etika, dan interpretasi ilmiah. **Sumber:** Diadaptasi dari ASA Statement, SAMPL Guidelines, CONSORT 2025, STROBE, PRISMA 2020, TRIPOD+AI, ICH E9(R1), dan FAIR Principles.

Review Questions

1. **Mengapa biostatistik tidak boleh dipahami hanya sebagai pencarian nilai p?**
   Jawaban: Karena biostatistik mencakup perumusan pertanyaan, desain studi, pengelolaan data, pemilihan analisis, estimasi efek, ketidakpastian, interpretasi, dan pelaporan ilmiah.
2. **Apa fungsi rencana analisis statistik?**
   Jawaban: Rencana analisis statistik menentukan outcome, variabel, metode analisis, asumsi, missing data, analisis sensitivitas, dan pelaporan sebelum hasil dilihat.
3. **Apa yang dimaksud estimand dalam ICH E9(R1)?**
   Jawaban: Estimand adalah definisi presisi tentang efek terapi atau parameter klinis yang ingin diestimasi dalam penelitian.
4. **Mengapa nilai p tidak cukup dilaporkan sendirian?**
   Jawaban: Nilai p tidak menunjukkan besar efek, relevansi klinis, atau kebenaran hipotesis, sehingga perlu disertai effect size dan confidence interval.
5. **Apa fungsi confidence interval?**
   Jawaban: Confidence interval menunjukkan rentang estimasi yang kompatibel dengan data dan membantu menilai presisi serta ketidakpastian hasil.
6. **Apa fungsi FAIR Principles?**
   Jawaban: FAIR Principles membantu memastikan data ilmiah dapat ditemukan, diakses, diintegrasikan, dan digunakan ulang secara bertanggung jawab.
7. **Pedoman apa yang digunakan untuk randomized trials?**
   Jawaban: Randomized trials sebaiknya dilaporkan menggunakan CONSORT.
8. **Pedoman apa yang digunakan untuk systematic review dan meta-analysis?**
   Jawaban: Systematic review dan meta-analysis sebaiknya dilaporkan menggunakan PRISMA.
9. **Pedoman apa yang digunakan untuk model prediksi klinis berbasis regresi atau machine learning?**
   Jawaban: Model prediksi klinis sebaiknya dilaporkan menggunakan TRIPOD+AI.
10. **Apa pesan akhir dari pembelajaran biostatistik terapan?**
    Jawaban: Pesan akhirnya adalah bahwa statistik harus digunakan secara terencana, transparan, etis, dapat direproduksi, dan ditafsirkan berdasarkan desain penelitian serta konteks ilmiah.

DAFTAR PUSTAKA

American Statistical Association. (2016). *Statement on statistical significance and p-values*. American Statistical Association.

Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., Ghassemi, M., Liu, X., Reitsma, J. B., van Smeden, M., Boulesteix, A. L., Camaradou, J. C., Celi, L. A., Denaxas, S., Denniston, A. K., Glocker, B., Golub, R. M., Harvey, H., Heinze, G., *et al*. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ, 385*, e078378.

Hopewell, S., Chan, A.-W., Collins, G. S., Hróbjartsson, A., Moher, D., Schulz, K. F., *et al*. (2025). CONSORT 2025 statement: Updated guideline for reporting randomised trials. *The Lancet*.

International Council for Harmonisation. (2019/2020). *ICH E9(R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials*.

Lang, T. A., & Altman, D. G. (2013). Basic statistical reporting for articles published in biomedical journals: The SAMPL guidelines. In *Science Editors' Handbook*. European Association of Science Editors.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., *et al*. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71.

STROBE Initiative. (n.d.). *STROBE checklists: Cohort, case-control, and cross-sectional studies*.

TRIPOD Statement. (2024). *TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods*.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., *et al*. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data, 3*, 160018.

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p-values: Context, process, and purpose. *The American Statistician, 70*(2), 129–133.

BIODATA PENULIS



Assoc.Prof. Dr.Arif Rachman,drg.,SH.,MH.,MM.,MTr.Hanla., Sp.Pros.,CIQnR.,CIQaR.,CSCBM.,CHRMP., CRMS.,CAIA.,FIMMA
Dosen Pendidikan Sarjana Kedokteran
Fakultas Kedokteran Militer Universitas Pertahanan RI

Penulis lahir pada tanggal 07 Februari 1971 di Surabaya – Jawa Timur, dari ayah yang bernama Dulmanan dan Ibu Romlah. Alamat di Jl. Cipinang Cempedak IV No. 60B Jakarta Timur 13340, e-mail ar604299@gmail.com / arifrachmanlecturer@gmail.com / arif.rachman@idu.ac.id

Pendidikan yang pernah ditempuh :Program Sarjana Jurusan Sarjana Pendidikan dokter gigi di FKG Unair Lulus Tahun 1995, Program Profesi dokter gigi di FKG Unair Lulus Tahun 1997, Program Spesialis Prosthodontia FKG Unair Lulus Tahun 2007, Program Magister Manajemen di Ubhara Surabaya Lulus Tahun 2008, Program Sarjana Ilmu Hukum di Ubhara Surabaya Lulus Tahun 2020, Program Magister Hukum di Ubhara Surabaya Lulus Tahun 2019, Program Doktoral Ilmu Kedokteran Dasar di Fakultas Kedokteran Unair Lulus Tahun 2019, Sertifikasi Peneliti Kuantitatif dan kualitataif di Quantum Lulus Tahun 2021.

Pekerjaan/Jabatan yang di pegang sekarang adalah: Dosen Tetap Fakultas Kedokteran Militer Universitas Pertahanan RI di Bogor, Kepala Laboratorium Fakultas Kedokteran Militer Universitas Pertahanan RI di Bogor, Dosen Tidak Tetap Universitas Adhirajasa Reswara Sanjaya di Bandung.

Organisasi Dosen: PDGI (anggota), IPROSI (anggota), MHKI (anggota), APEPINDO (anggota)

Karya Tulis dalam bentuk buku dan jurnal penelitian adalah:

1) Filsafat Statistika. Get Press Indonesia, ISBN: 978-624-255-436-4, April 2026
2) hADMSC-YTZP seeding improves peri-implant bone repair through increased expression of BMP-2, ALP, and OPG. Scientific Reports, 15(37800), pp. 1–12. Scientific Reports, 15(37800), pp. 1–12. www.nature.com/scientificreports https://doi.org/10.1038/s41598-025-21655-6
3) Biostatistik Kesehatan. Get Press Indonesia, ISBN: 978-643-255-030-4, September 2025
4) Reduction of Osx expression in the osseointegration process of dental implants with human adipose-derived mesenchymal stem cell intervention: an experimental study. Padjadjaran Journal of Dentistry, Volume 37, Number 2, July 2025; 142-154. https://doi.org/10.24198/pjd.vol37no2.62084
5) Dasar-Dasar Biologi Sel Dan Molekuler. Get Press Indonesia, ISBN: 978-623-125-848-9, Juni 2025
6) Sistem Informasi E-Hospital. Lakeisha, ISBN: 978-623-119-614-9, Maret 2025
7) Etika, Hukum, Dan Regulasi Dental Implan Untuk Drg Umum.Get Press Indonesia, ISBN: 978-623-125-659-1, Maret 2025
8) Pengantar Metodologi Kuantitatif Dan Kualitatif. CV. Saba Jaya Publisher ISBN 978-623-8730-33-9, Maret 2025
9) Prevention of Military Stress Fractures using the hadMCS Strategy Approach in Health Resilience Projections in National Defence: Preliminary Research. PJLSS 2024 Posted: Oct 19, 2024. Volume 22(2). 2024; 8791-8798. https://doi.org/10.57239/PJLSS-2024-22.2.00663 (Q4).
10) The Correlation of Exclusive and Non-Exclusive Asi on Stunting In The Context of State Defense: Initial Research on PCBM In Bogor District. PJLSS 2024 Posted: Oct 20, 2024. Volume 22(2). 2024; 8976-8985. https://doi.org/10.57239/PJLSS-2024-22.2.00679 (Q4).
11) Acceleration of Healing of TB Patients through Dietary Supplements on Anti-Tuberculosis Drugs. PJLSS 2024 Posted: Oct 19, 2024. Volume 22(2). 2024; 8799-8807. https://doi.org/10.57239/PJLSS-2024-22.2.00664 (Q4).
12) Prevention Of Stunting In The First 450 Days Of Life (FDL) Through Fortification Of Tuna Fish Cake And Formula Programming In Supporting Indonesia's National Defense In The Health Sector.

PJLSS 2024 Posted: Oct 20, 2024. Volume 22(2). 2024; 8986-8996. https://doi.org/10.57239/PJLSS-2024-22.2.00680. Q4)

13) Increased Runx-2 Expression In The Y-TZP Biomaterial Osteointegration Process With hADMSC Seeding. SEEJPH 2024 Posted: 02-08-2024. Volume XXV 2024; 216-224.https://doi.org/10.70135/seejph.vi.1470 (Q4).
14) Manajemen Epidemiologi, Get Press Indonesia, ISBN: 978-623-125-387-3, September, 2024.
15) Pengantar Hukum Indonesia, Get Press Indonesia, ISBN: 978-623-125-239-5, Juni 2024.
16) Komunikasi Dalam Praktik Keperawatan (Komunikasi Efektif Layanan Kesehatan), Agdosi Makassar, ISBN: 978-623-10-0651-6, Mei 2024
17) Manajemen Risiko dan Asuransi, PT. MAFY Literasi Indonesia, ISBN: 978-623-863-871-0, Mei 2024.
18) Metode Penelitian dalam Pendidikan_2, Get Press Indonesia, ISBN: 978-623-125-155-8, April 2024.
19) Dasar Biostatistika untuk Peneliti, Get Press Indonesia, ISBN: 978-623-125-137-4, April 2024.
20) Pengantar Manajemen, Get Press Indonesia, ISBN: 978-623-125-138-1, April 2024.
21) Konsep Autoimun, PT. MAFY Literasi Indonesia, ISBN: 978-623-8575-21-3, Maret 2024.
22) An Innovative Approach to Dental Implantology Through DPMSC Therapy in Military Dentistry, Tahta Media Group, ISBN: 978-623-147-325-7, Februari 2024.
23) Metodologi Penelitian Kuantitatif, Kualitatif, dan R&D, CV. Saba Jaya Publisher ISSN: 978-623-09-7582-0, Januari 2024
24) Fakultas Kedokteran Militer Universitas Pertahanan Cita dan Jejak Sejarah, Unhan Press,Januari Tahun 2021.
25) Bioetika Multidisplin, Multidisiciplinary Prespectives on Bioethics, Unhan Press, Januari Tahun 2022.
26) Aesthethic and masticatory rehabilitation on post mandibular resection with combination of hollow obturator and hybrid prosthesis, Dent. J. (Maj. Ked. Gigi), Vol. 42. No. 2 April–June 2009: 76–81
27) Biocompatibility of Yttria-Tetragonal Zirconia Polycrystal Seeded with Human Adipose Derived Mesenchymal Stem Cell, ACTA INFORM MED. 2018 DEC; 26(4): 249-253 (Q3)

28) Addition Of Adipose Derived Stem Cell To Beta Tricalcium Phosphate and Human Cancellous Bone for Craniofacial Bone Tissue Engineering: An In Vitro Study, Malaysian Journal of Medicine and Health Sciences. 2020; 16(4): 83-86 (Q3)
29) Comparison of Bone regeneration in hADMSC Versus hUCBMSC with hBMMSC as a Reference: A Literature Review of Potential Bone Regeneration, Research J. Pharm. and Tech. 14(4): April 2021 (Q3)
30) CBCT: Tracing Dominant Region of Periodontitis on Psychological Stress, Sys Rev Pharm. 2020;11(11):1567-1572 (Q2)
31) Ideal Bone Defects Distance on Orthodontic Tooth Movement for Preparation of hADMSC-Scaffold Chitosan Intervention, Sys Rev Pharm. 2020;11(8):571-575 (Q2)
32) Accuracy of Volume of Interest Determination with Cone-Beam Computed Tomography in Periodontitis due to Psychological Stress., Journal of International Dental and Medical Research ISSN 1309-100X http://www.jidmr.com (Q3).
33) Antitumor Effects of Myrmecodia pendans: a Scoping Review, Journal of International Dental and Medical Research, Volume 16 Number 2 , 2023;1–5, ISSN 1309-100X (Q3).
34) Evaluation of the Use of Antibiotics on Therapy Results of Sepsis Patients in the Intensive Care Unit (ICU) of Fatmawati Hospital,Jakarta, J Med Chem Sci. 2023;7(1):262–74. (Q3).
35) Enhancing Denture Care Efficiency: Mobile Prosto Open-Source Software for Indonesian National Army Soldiers, J Int Dent Med Res 2024; 17(1): 209-214); ISSN 1309-100X. (Q3).
36) The Impact of Lifestyle Factors on the Human Microbiome, Medical Studies and Health Journal (SEHAT) Vol 1 (1) 2024 : 113-122.
37) Exploring the Therapeutic Potential of Ant's Nest Plant Extract on Epithelial Dysplasia in an Oral Squamous Cell Carcinoma Model: A Study on Apoptosis Markers in Sprague Dawley Mice, Journal of International Dental and Medical Research, Volume 17 Number 2, 2024;543–548, ISSN 1309-100X (Q3).
38) Pendidikan Kesehatan, Pemeriksaan dan Deteksi Dini Pencegahan Penyakit Tidak Menular melalui Kegiatan Olah Raga Fisik di Wilayah Kerja Puskesmas Tamalanrea Makassar, Sahabat Sosial Jurnal Pengabdian Masyarakat, e-ISSN: 2964-9196, Vol.2 No.4 September 2024

BIODATA PENULIS



Dr. apt. Drs. Timbul Partogi Haposan Simorangkir, M.Si.C I T.
Dosen Fakultas Farmasi Universitas Pertahanan

Penulis Lahir pada tanggal 16 Nopember 1968 di Surabaya, dari Ayah : Kapten Pelaut ALM. Marulam Simorangkir dan Ibu.: Alm. Anita Hinsa Sitompul. Alamat di Jl. Gudang Utara 14 Sumur Bandung, Kodya Bandung ,Email: tphsimorangkirsimon4@gmail. Pendidikan yang pernah ditempuh: Program Sarjana Farmasi Unair Lulus Tahun 1992, Program Profesi Farmasi Unair Lulus Tahun 1993, S2 bidang Farmasetika/Teknologi Farmasi ITB bandung Lulus Tahun 2002, S3bidang Farmasetika/Teknologi Farmasi ITB bandung Lulus Tahun 2015. Penulis pernah melakukan ko riset dengan Tim Prof Uekusa dari TIT Jepang pada THN 2013 dan melaksanakan riset bersama dengan BRIN THN 2019.

Pengalaman Mengajar

No	Institusi	Jabatan	Periode Kerja
1.	Akademi Keperawatan Dustira Cimahi	Dosen Tetap Ilmu Farmasi	2003-2013
2. 3	Akademi Fisioterapi Dustira Cimahi	Dosen Tetap Ilmu Farmasi	2003 – 2013

4.	Sekolah Tinggi Farmasi Bandung	Dosen Tetap Ilmu Farmasi	2013-2019
	Universitas Bhakti Kencana Bandung		2010-2013
5.	Universitas Setia Budi Sol	Dosen LB Ilmu Farmasi	2013-2020
6.	Universitas Halu Oleo Kendari	Dosen LB Ilmu Farmasi	2021-2022
7.	Unjani Cimahi	Dosen LB Ilmu farmasi	2022-2023
8.	Univ Pertahanan RI	Dosen LB Ilmu Farmasi	2020-sekarang
		Dosen Tetap Ilmu Frmasi	

Karya Tulis dalam bentuk buku dan jurnal penelitian adalah:

1. Nadya F., Dolih G., Yasmiwar S. TPH Simorangkir, *Formulasi Tablet Hisap yang Mengandung Ekstrak Akar Ginseng Korea (Panax ginseng C.A. Meyer) dan Ekstrak Rimpang Temulawak (Curcuma xanthorrhiza Roxb.)*, Skripsi Unpad 2006.
2. Tenny., M.Abdassah, TPH Simorangkir, *Formulasi Granul yang Mengandung Ekstrak Akar Ginseng Korea (Panax ginseng C.A. Meyer) dan Ekstrak Rimpang Temulawak (Curcuma xanthorrhiza Roxb.)*, Skripsi Unpad 2006.
3. Identification of Physical Interaction between Antimalarial Drugs Combination Artesunate-Amodiaquine hydrochloride, International Journal of Pharmacy and Pharmaceutical Sciences Academic Sciences 5(3), 206-210.
4. Influence of Milling Process on Solid States Transformation of Artesunate. The 3rd International Conference on Pharmacy and Advanced Pharmaceutical Sciences Proceedings Book (1), 39-46.
5. Preparasi dan Karakterisasi Polimorfisme Obat Antimalaria Artesunate, Jurnal Sains Materi Indonesia 15(2), 88-93

6. Study of Artesunate (Antimalarial Drug) Polymorphism. The 24th International FAPA Congress, Nusa Dua Bali , September 13 – 16 th, 2012.
7. Aplikasi Difraksi Sinar-X Metode Serbuk untuk Identifikasi dan Karakterisasi Polimorfisme Obat Anti Malaria Artesunate. Seminar Nasional dan Workshop Hamburan Neutron dan Sinar-X ke 9 . Batan, Serpong, Banten 6 November 2013.
8. Diktat Pengantar Industri Farmasi dan Pengenalan CPOB untuk Tenaga Teknis Kefarmasian dan Tenaga Kefarmasian 2014
9. Pengaruh Penambahan L-arginin terhadap kelarutan Hidroklortiazid (HCT) 2018
10. Pengaruh Penambahab Benzamide terhadap kelarutan Indapamide 2019
11. Diktat Farmasetika Dasar untuk Mahasiswa Sarjana Farmasi 2022
12. The Potential of Indonesian Natural Materials as Immunomodulators and Tonics for National Resilience of Public Health in the Era Covid-19 Pandemic 2022
13. Preparation and physicochemical characterization of p-methoxycinnamic acid - Succinic acid Co Crystal by Solvent evaporation technique 2022
14. Fixed-Dose Combination (FDC) Formulation of Quinine Sulphate-Doxycycline (Qidox) for Malaria Therapy 2023
15. Developement and Validation of method for analysis fafipiravir and remdesivir in volumetric absorptive micro sampling with HPL tandem Mass Spectrophotometer 2023
16. The Role of Reserve Components in Supporting Indonesia's Defense Forces to Face Non-Military Threats 2023
17. Faktor-Faktor yang Berhubungan dengan Komplikasi Apendisitis dan Efektivitas Penggunaan Antibiotik Profilaksis pada Pasien Apendektomi di RSPAD Gatot Soebroto 2023
18. Lipid droplets in endosymbiotic Symbiodiniaceae spp. associated with corals 2024
19. Herbal Extract of Red Dragon Fruit as a Basic Ingredient of Disclosing Agent that Has a Role as an Antioxidant, Antibacterial, and Antibiofilm Against Streptococcus mutans in Caries Prevention in the Field of Dentistry 2024
20. Sistem Informasi E Hospital Penerbit Lakesha 2025
21. Polifenol : Struktur, Sumber dan Manfaat Kesehatan, Penerbit CV Patra Media Grafindo Bandung 2025