

# Index

- absolute benefit increase (ABI) 452–3
- absolute risk reduction (ARR) 452, 454
- activities of daily living (ADL)
  - Barthel Index 6, 29–30, 507
  - cross-sectional analysis 286–7
  - item response theory 162
  - questionnaire design 61
  - rehabilitation 12
- acupuncture 466–7
- acute myocardial infarction 12
- ADF *see* asymptotically distribution-free (ADF) factor analysis
- ADL *see* activities of daily living (ADL)
- age adjustments 435–8
- AIMS *see* Arthritis Impact Measurement Scale (AIMS)
- analysis of variance (ANOVA)
  - cross-sectional analysis 291–7
  - distribution 295
  - factor analysis 134
  - longitudinal data 329, 335–8
  - models 296–7
  - multi-item scales 125–6
  - multivariate 351–2, 391
  - regression 291–4, 296–7
  - repeated measures 335–8
  - reporting issues 388, 391
- ANOVA *see* analysis of variance (ANOVA)
- anxiety *see* Hospital Anxiety and Depression Scale (HADS); Pain Anxiety Symptoms Scale (PASS)
- Apgar scores 44, 129
- AQLQ *see* Asthma Quality of Life Questionnaire (AQLQ)
- area under the curve (AUC)
  - longitudinal data 305–8, 324–6, 352–3
  - missing data 357
  - quality-adjusted survival 411–12
  - reporting issues 390
- Aristotle 5
- ARR *see* absolute risk reduction (ARR)
- Arthritis Impact Measurement Scale (AIMS) 449
- association of variables
  - graphical methods 299–301
- Asthma Quality of Life Questionnaire (AQLQ)
  - clinical interpretation 442–3, 452–3
  - responsiveness 106
  - see also* Paediatric Asthma Quality of Life Questionnaire (PAQLQ)
- attrition 268
  - bias 461–2
- AUC *see* area under the curve (AUC)
- auto-correlation
  - error structure 333–5, 344
  - exchangeable 335
  - longitudinal data 329, 330–5, 344, 350
  - matrices 330–3
  - patterns 333–5
  - structure selection 350
- auto-regression 338–40
- available-case analysis 356–7
- bar charts 298, 300, 391–3
- Barthel Index of Disability (BI) 6, 29–30, 507
- baselines
  - assessment 236
  - cross-sectional analysis 290–1
  - variables 394
- battery scores 37
- BI *see* Barthel Index of Disability (BI)
- biases
  - clinical interpretation 438–41
  - differential item functioning 177, 178
  - meta-analysis 460–2, 474–5
  - missing data 355, 360–2, 381, 383
  - response shifts 438–41
  - selective reporting 438–41
- binary data
  - kappa coefficient 92–3
  - meta-analysis 467–8, 470

- binary data (*Continued*)
  - proportion of agreement 92
  - reliability 92–3
- binomial proportions 276–7
- BMJ* *see* *British Medical Journal (BMJ)*
  - checklists
- bone marrow transplants (BMT) 212
- Bonferroni corrections 266–7
- bootstrap methods
  - cross-sectional analysis 284
  - quality-adjusted survival 418
- box-and-whisker plots 300–2, 313–15
- BPI *see* Brief Pain Inventory (BPI)
- BPRS *see* Brief Psychiatric Rating Scale (BPRS)
- brachytherapy 346
- breast cancer
  - missing data 359, 362
  - quality-adjusted survival 412–18, 420–5
- Brief Pain Inventory (BPI) 443
- Brief Psychiatric Rating Scale (BPRS) 333, 349
- British Medical Journal (BMJ)* checklists 385
- cancer *see* FACT-G; Functional Living Index – Cancer (FLIC); malignancies; QLQ-C30
- case-control pairs 254–5, 257
- CAT *see* computer-adaptive testing (CAT)
- categorical data
  - cross-sectional analysis 277–9
  - factor analysis 150–1
  - see also* ordered categorical data
- causal
  - indicators 34, 156
  - items 67, 71–2, 127–9
  - models 155–8, 159
  - variables 7
    - item response theory 168, 188
    - measurement scales 33–5, 45, 46–7
    - multi-item scales 110
    - weighted sum scores 218
- ceiling effects
  - differential item functioning 186
  - longitudinal data 307
- questionnaire design 71–2
- censoring, informative 368
- CF *see* cognitive functioning (CF) scale
- CFA *see* confirmatory factor analysis (CFA)
- chi-squared tests
  - cross-sectional analysis 277–9
  - differential item functioning 187
  - reporting issues 388
  - sample sizes 255–6
  - statistical tables 513
- children, instruments 211
- chronic illness 5, 24–5, 196, 493–4
- chronic lymphatic leukaemia (CLL) 212
- Chronic Respiratory Questionnaire 441–2
- CI *see* confidence intervals (CI)
- clinical interpretation 427–55
  - absolute levels 429–30
  - adjustments for age and gender 435–8
  - biases 438–41
  - effect size 427, 446–51
  - guidelines 214
  - impact of state of QoL 443–4
  - life events 444–6
  - literature 455
  - meta-analysis 450–1
  - minimal clinically important difference 441–3
  - number needed to treat 452–4
  - patient-orientated methods 427, 446–54
  - patient variability 451–2
  - percentages 430–1
  - population norms 427, 431–41
  - response shifts 438–41
  - selective reporting 438–41
  - significance tests 428–9
  - standardised response mean 447–8
  - threshold values 430–1
  - time variation 429–30
- clinical significance 391
- clinical trials *see* randomised clinical trials (RCTs)
- clinimetrics
  - measurement scales 44–5
  - multi-item scales 110, 128
  - questionnaire design 72–3

- CLL *see* chronic lymphatic leukaemia (CLL)
- Cochrane Handbook for Systematic Reviews* 457, 466
- Cohen's effect size index 447
- complete-case analysis 356
- compliance
  - acceptable levels 229–30
  - causes/consequences of non-compliance 227–9
  - improving 229, 234
  - measuring 226–7
  - meta-analysis 461–2
  - missing data 383
  - protocols 234
  - randomised clinical trials 223–4, 225–31, 234
  - recording 230
  - reporting issues 388, 400
  - schedule and attrition 322–4
- computer-adaptive testing (CAT) 189–205
  - advantages/disadvantages 199
  - cost and practicalities 203
  - differential item functioning 204
  - item banks 191–2
  - item calibration 191, 193
  - item linking 194–5
  - multidimensional 203
  - simulations 200–1
  - software 201–2
  - stopping rules 200–1
  - test equating/linking 191, 195–6
  - test information functions 191, 196–8
- computer-assisted testing 204
- confidence intervals (CI)
  - cross-sectional analysis 276–7, 281
  - meta-analysis 463, 472
  - missing data 371
  - multi-item scales 113–14, 115
  - reporting issues 385, 389, 394
  - variability 319
- confidentiality protocols 240–1
- confirmatory factor analysis (CFA) 131, 141, 158–60
  - dimensionality 193
- conjoint analysis 408–9
- consistency *see* reliability
- CONSORT statement 247, 385, 401
- construct validity
  - convergent validity 85–6, 87
  - dimensionality 83
  - discriminant validity 86–7, 88
  - homogeneity 83
  - known-groups validation 84–5
  - multi-item scales 109, 116–23
  - multitrait–multimethod analysis 87–91
  - overlap between latent variables 83
  - questionnaire design 51
  - single-item scales 77–8, 83–91
  - summated ratings 77–8, 83–91
- constructs, definition 32–3
- content validity 64, 77–80
- continuous data
  - meta-analysis 464–7, 469–70
  - numerical 275
- convergent validity 77–8, 85–7, 109, 117–19
- correlations
  - confidence intervals 113–14
  - factor analysis 132–4, 154–5
  - multi-item scales 111–16
  - overlap correction 116
  - range of variables 112
  - rank correlations 115
  - significance tests 112–13, 114–15
- cost-utility ratios 413–14
- covariance
  - factor analysis 133–4
  - repeated measures 348–50
- criterion validity 77–8, 81–3
- Cronbach's  $\alpha$ 
  - factor analysis 147
  - item response theory 174–5
  - measurement scales 38, 39
  - multi-item scales 123–9
  - questionnaire design 72
- cross-sectional analysis
  - aggregated continuous data 275–6
  - analysis of variance 291–7
  - association of variables 299–301
  - baselines 290–1
  - binomial proportions 276–7
  - categorical data 277–9
  - chi-squared tests 277–9
  - clinical interpretation 429

- cross-sectional analysis (*Continued*)  
 graphical methods 297–303  
 linear regression 285–8  
 logistic regression 288–90  
 nominal data 273–4  
 non-Normal distributions 275, 282–4,  
 288–90, 295  
 Normal distributions 275, 280–1,  
 285–8, 295  
 numerical data 275  
 ordered categorical data 274–5  
 patient profiles 301–2  
 two-group comparisons 276–85
- cultural differences  
 differential item functioning 184–5  
 instruments 214  
 questionnaire design 70, 74
- data-derived effect sizes 427
- D-CAT *see* depression -  
 computer-adaptive testing (D-CAT)
- DCEs *see* discrete choice experiments  
 (DCEs)
- decision-making 15–16
- definitions of QoL 4–5, 52
- depression - computer-adaptive testing  
 (D-CAT) 201
- depression *see* Hamilton Rating Scale for  
 Depression (HRSD); Hospital  
 Anxiety and Depression Scale  
 (HADS)
- detection bias 461
- DIF *see* differential item functioning (DIF)
- differential item functioning (DIF) 74,  
 176–88  
 ceiling/floor effects 186  
 computer-adaptive testing 204  
 item bias 177, 178  
 item response theory 179, 204  
 literature and software 187  
 logistic regression 181–5, 186  
 Mantel–Haenszel test 179–81, 187  
 pseudo DIF 183–4  
 quantification 185  
 tips 185–7  
 translations 183, 185  
 uniform/non-uniform 182, 186
- difficulty, item 164–5, 168, 202–3
- dimensionality  
 construct validity 83  
 item response theory 167–8, 193, 203  
 multi-item scales 109, 122–3
- direct standardisation 435
- disability  
 Barthel Index 6, 29–30, 507  
 questionnaire design 73
- discrete choice experiments (DCEs)  
 408–9
- discriminant validity 77–8, 86–8, 109,  
 117–19
- discrimination, item 164–6, 173, 175–6,  
 196
- discriminative  
 instruments 101  
 scales 46
- disease-specific instruments 23–7,  
 210–11, 489–502  
 PAQLQ 26–7, 98–9, 501–2  
 QOLIE-89 26, 497–500  
 Rotterdam Symptom Checklist  
 25–6, 153, 155–7, 176–7,  
 363, 495–6  
*see also* FACT-G; QLQ-C30
- domain-specific instruments 27–30,  
 503–6  
 McGill Pain Questionnaire 28–9,  
 88–9, 504  
 MFI-20 11–12, 29, 68, 505–6  
*see also* Hospital Anxiety and  
 Depression Scale (HADS)
- dummy variables  
 ANOVA 294  
 missing data 374
- EFA *see* exploratory factor analysis (EFA)
- effect indicators 34, 156
- effect size (ES)  
 clinical interpretation 427, 446–51  
 Cohen's effect size index 446–7  
 effect size statistic 448–51  
 Glass's delta 448  
 meta-analysis 450–1  
 sample sizes 248–52  
 sensitivity and responsiveness 102,  
 105–6  
 standardised response mean 447–8

- eigenvalues 136–7, 145–6  
 EM algorithm 380  
 endpoints 303, 399  
 EORTC *see* QLQ-C30  
 epilepsy 26, 57, 80, 497–500  
 Epilepsy Surgery Inventory 55 (ESI-55)  
   80  
 EPO *see* erythropoietin (EPO)  
 EQ-5D *see* EuroQol  
 equivalence trials 224, 263–5  
 equivalent-forms reliability 91, 101  
 erythropoietin (EPO) 210  
 ES *see* effect size (ES)  
 ESI-55 80  
 essentially tau-equivalent tests 42  
 EuroQol 6, 22, 485–6  
   instrument selection 210  
   quality-adjusted survival 410–11  
 exchangeable auto-correlation 335  
 expectations model 7  
 explained variance 136–7  
 exploratory factor analysis (EFA) 131,  
   141, 148, 155–6, 158–60
- face validity 64, 73, 79–80  
 FACIT 24–5, 192, 493–4  
 FACT-G 24–5, 493–4  
   instrument selection 212  
   longitudinal data 310–11  
   questionnaire design 51, 63, 75
- factor analysis  
   applications 141–3  
   assumptions 149–53  
   asymptotically distribution-free 152  
   categorical data 150–1  
   causal models 155–8, 159  
   choices and decisions 143–9  
   confirmatory 131, 141, 158–60  
   correlations 132–4, 154–5  
   dimensionality 193  
   eigenvalues 136–7, 145–6  
   exploratory 136, 141, 148, 155–6,  
     158–60  
   factor loadings 138  
   HADS questionnaire 132–52  
   historical development 141  
   literature 160  
   maximum likelihood estimation 146–7
- multi-item scales 122–3  
 number of factors 145–6  
 oblique axes 148–9  
 path diagrams 134–5  
 rotation 138–40, 147–9  
 sample size 144–5, 154  
 scale validation 142  
 structural equation modelling 131,  
   158–60
- factor loadings 138  
 FAQ *see* Functional Activity  
   Questionnaire (FAQ)  
 fatigue  
   cross-sectional analysis  
     281, 294  
   item response theory 199–200  
   longitudinal data 325–6  
   questionnaire design 68  
   *see also* MFI-20; Nordic Myeloma  
     Study Group
- feasibility 214  
*F*-distribution  
   reporting issues 388  
   sensitivity 102–3, 105  
   statistical tables 514
- field-testing 69–74  
   cultural differences 70, 74
- Fisher's exact test 278  
 Fisher's *F*-test 291, 294  
 five-and-twenty rule 230, 462  
 fixed-effects models 344–6  
   *see also* repeated measures
- FLIC *see* Functional Living Index –  
   Cancer (FLIC)
- floor effects  
   differential item functioning 186  
   longitudinal data 307  
   questionnaire design 71–2
- focus groups 56–7  
 forest plots 471–2
- Functional Activity Questionnaire  
   (FAQ) 30
- Functional Assessment of Cancer  
   Therapy – General (FACT-G) 24–5,  
   493–4
- Functional Assessment of Chronic Illness  
   Therapy (FACIT) 24–5, 192

- Functional Living Index – Cancer (FLIC) 14  
 construct validity 89–91  
 missing data 369  
 multi-item scales 114, 115
- funnel plots 474–5
- GEE *see* generalised estimating equations (GEE)
- gender adjustments 435–8
- General Perceived Health (GPH) scale 444–5
- generalised estimating equations (GEE)  
 longitudinal data 329, 342–4  
 reporting issues 390–1
- generic instruments 19–23, 210–11, 479–87  
 EuroQol 6, 22, 210, 411, 485–6  
 Nottingham Health Profile 6, 20–1, 107, 228–9, 481  
 Patient Generated Index 7, 22–3, 487  
*see also* SF-36; Sickness Impact Profile (SIP)
- global  
 questions 35–6, 47  
 rating scales 442  
 scores 163
- goodness-of-fit indexes 169, 187, 188
- GPH *see* General Perceived Health (GPH) scale
- graphical methods  
 association of variables 299–300  
 cross-sectional analysis 297–302  
 longitudinal data 305–20  
 reporting issues 391–4, 396–7  
*see also* area under the curve (AUC)
- Guttman scales  
 item response theory 174–5  
 questionnaire design 61–2
- HADS *see* Hospital Anxiety and Depression Scale (HADS)
- Hamilton Rating Scale for Depression (HRSD) 463, 470–1, 475–6
- HDS *see* Hodgkin’s disease survivors (HDS)
- Headache Impact Test (HIT-6) 182, 192, 197–8
- health-economics scores 221–2, 224
- health outcomes assessments 3
- health-related quality of life (HRQoL) 3, 4
- Health Utilities Index (HUI) 283, 410
- healthy-years equivalent (HYE) 424–5
- hepatitis C 10
- heterogeneity  
 factor analysis 154  
 meta-analysis 472–3
- hierarchical scales 366
- historical development of QoL 5–8
- HIT-6 *see* Headache Impact Test (HIT-6)
- HIV Overview of Problems – Evaluation System (HOPES) 84–5, 86–7, 104, 294
- Hodgkin’s disease survivors (HDS) 288
- homeopathy 339
- homogeneity 83, 279
- HOPES 84–5, 86–7, 104, 294
- horizontal mean imputation 372–3
- Hospital Anxiety and Depression Scale (HADS) 503  
 clinical interpretation 440  
 criterion validity 82  
 cross-sectional analysis 278–9, 282–4, 289–90  
 factor analysis 132–53  
 instruments 27–8  
 longitudinal data 312–19, 351, 354  
 measurement scales 35, 38  
 multi-item scales 109  
 sample sizes 250–4, 258–61, 265–6  
 validation 82
- hot deck imputation 379–80, 383
- HRQoL *see* health-related quality of life (HRQoL)
- HRSD *see* Hamilton Rating Scale for Depression (HRSD)
- HUI *see* Health Utilities Index (HUI)
- HYE *see* healthy-years equivalent (HYE)
- Hypericum 463, 470, 475
- hypertension 9, 18–19
- IADL *see* instrumental activities of daily living (IADL)

- ICC *see* intra-class correlation coefficient (ICC); item characteristic curves (ICC)
- impact of state of QoL 443–4
- imputed data 355–8
- meta-analysis 462
  - missing forms 369–81, 383
  - missing items 364–8
- incomplete data 310–13, 320
- indicator variables 33–5, 45, 46–7
- indirect standardisation 435
- individual profiles 313
- information
- functions 173, 191
  - leaflets 241–2
- informative censoring 368
- instrumental activities of daily living (IADL) 30
- instruments 3, 19–30
- adding ad-hoc items 215
  - checklist 213–15
  - children 211
  - cultural differences 214
  - development 213
  - documentation 213
  - health-economics scores 221–2
  - interpretation guidelines 214–15
  - item response theory 220–1
  - literature 211, 221
  - meta-analysis 462
  - multi-item scales 215–22
  - norming systems 218–20
  - patient preferences 17–19
  - questionnaire design 51–76
  - reporting issues 386, 399
  - scoring procedure 214, 215–22
  - selection criteria 209–15
  - sensitivity and responsiveness 101
  - summated scales 216–17
  - target populations 213–14
  - treatment-specific 210–11
  - T-scores 218–20
  - utility-based 410–11
  - validation 213, 462
  - weighted sum scores 217–18
  - Z-scores 218–20
- see also* disease-specific instruments; domain-specific instruments; generic instruments
- inter-rater reliability 91, 99–100
- internal reliability 91
- multi-item scales 109, 123–8
- internal validity 39
- International Breast Cancer Study Group (IBCSG) 413–18, 420–5
- interpretation *see* clinical interpretation
- interval scales 59
- interviews
- patient 55
  - questionnaire design 54–6
  - specialist 54
- intra-class correlation coefficient (ICC) 96–7
- inverse variance weighted 469
- IRT *see* item response theory (IRT)
- item
- banks 191–2
  - calibration 191, 193
  - convergence 79
  - definition 31–2
  - difficulty 164–5, 168, 202–3
  - discrimination 164–5, 173, 175–6, 196
  - information functions 173
  - linking 194–5
  - non-response 358
  - reduction 72–3
  - see also* differential item functioning (DIF); item response theory (IRT)
- item characteristic curves (ICC) 163–5, 167, 175
- item-misfit indexes 169, 172
- item response theory (IRT) 43, 161–76, 187–8
- computer-adaptive testing 189–205
  - computer-assisted testing 204
  - cost and practicalities 203
  - differential item functioning 179, 204
  - difficulty 164–5, 168, 202–3
  - discrimination 164–6, 173, 175–6, 196
  - Guttman scales 174–5
  - instrument selection 220–1
  - item banks 191–2
  - item calibration 191, 193

- item response theory (IRT) (*Continued*)
  - item characteristic curves 163–5, 167, 175
  - item linking 194–5
  - literature 188, 204
  - logistic models 165–72, 175–6, 189–91
  - logistic regression 167
  - polytomous models 171, 175–6
  - questionnaire design 54, 61–2
  - simulations 200–1
  - software 187, 201–2
  - test design 173–4
  - test equating/linking 191, 195–6
  - test information functions 173, 191, 196–8
  - threshold parameters 175–6
- Kaplan–Meier survival curves
  - longitudinal data 322–3
  - missing data 359
  - quality-adjusted survival 415–16
  - sample sizes 261
- kappa coefficient 92–5
- Karnofsky Performance Scale 6
- known-groups validation 84–5
- LASA *see* linear analogue self-assessment (LASA)
- last value carried forward (LVCF) 370–1, 377–9
- latent roots *see* eigenvalues
- latent variables 32–3, 36–7, 73
  - factor analysis 136, 155
- Life Events Index 445–6
- Likert summated scales
  - item response theory 41
  - multi-item scales 110
  - questionnaire design 54, 59, 63
- line plots 393
- linear analogue self-assessment (LASA) 6, 60–1, 369
- linear regression
  - cross-sectional analysis 285–8, 296–7
  - longitudinal data 329, 338–40
- local independence 134, 168
- logarithmic transformations 266, 285
- logistic item response models 165–72, 175–6
  - assumptions 167–8
  - computer-adaptive testing 189–91
  - fitting the models 168–72
  - logistic regression 167
  - longitudinal data 351
  - polytomous models 171, 175–6
  - model 165–166
  - selection 171
- logistic regression
  - cross-sectional analysis 288–90
  - differential item functioning 181–4, 186
  - item response theory 167
- logits
  - cross-sectional analysis 288–90
  - differential item functioning 185
  - item response theory 166–7
  - meta-analysis 467
- longitudinal data 305–27
  - analysis of variance 329, 335–6
  - area under the curve 305–8, 324–6, 352–3
  - auto-correlation 329, 330–5, 344, 350
  - compliance with schedule and attrition 322–4
  - covariates 348–50
  - fixed-effects models 344–6
  - generalised estimating equations 329, 342–4
  - graphical methods 308–20
  - incomplete data 310–13, 320
  - individual profiles 313
  - interactions 340
  - linear regression 329, 338–40
  - logistic models 351
  - missing data 352
  - mixed-effects models 346
  - modelling 329–54
  - multilevel models 329, 335, 347
  - multivariate analysis of variance 351–2
  - paired data 341–2
  - percentages 308–9
  - random-effects models 344–6
  - repeated measures 329, 335–50
  - reporting 322–6

- tabular presentation 320–2
- treatment comparisons 324–6
- lung cancer
  - longitudinal data 308–9
  - randomised clinical trials 226, 233
- LVCF *see* last value carried forward (LVCF)
- McGill Pain Questionnaire (MPQ) 28–9, 88–9, 504
- McNemar test 257
- manifest variables 33
- Mann–Whitney *U*-test
  - cross-sectional analysis 283
  - quality-adjusted survival 418
  - sample sizes 258
- MANOVA *see* multivariate analysis of variance (MANOVA)
- Mantel–Haenszel test 179–81, 187
- MAR *see* missing at random (MAR)
- Markov chain imputation 375–9, 383
- maximum allowable difference 263
- maximum likelihood estimates (MLEs)
  - factor analysis 146–7
  - meta-analysis 469
- MCAR *see* missing completely at random (MCAR)
- MD *see* raw mean difference (MD)
- mean and sigma method 195
- measurement scales 3, 5, 31–48
  - causal variables 7, 33–5, 45, 46–7
  - clinimetrics 44–5
  - constructs 32–3
  - discriminative 46
  - evaluative 46
  - factor analysis 42–3, 142
  - global questions 35–6, 47
  - indicator variables 33–5, 45, 46–7
  - item response theory 43
  - latent variables 32–3
  - necessary causes 45–6
  - parallel tests 41–2, 47
  - precision 38–9
  - predictive 46
  - psychometrics 40–5, 47
  - questionnaire design 58–63
  - reliability 38, 46
  - scale scores 19, 143
  - scope 40
  - standardisation 191
  - sufficient causes 45–6
  - validation 39–40, 142
  - see also* multi-item scales; single-item scales
- measures of health status *see* generic instruments
- medical decision-making 15–16
- Medical Outcomes Study (MOS) 10
  - see also* SF-36
- MEDLINE 54
- meta-analysis 457–75
  - binary data 467–8, 470
  - combined studies 468–71
  - continuous data 464–7, 468–71
  - effect size 450–1
  - forest plots 471–2
  - funnel plots 474–5
  - heterogeneity 472–3
  - instrument validation 462
  - literature 457, 459, 475
  - measures of treatment effect 464–8
  - ordinal data 468, 471
  - publication bias 460–2, 474–5
  - quality assessment 459–62
  - summarising results 463–4
- methodology sections 398
- MFI-20 11–12, 29, 68, 505–6
- MIMIC *see* multiple-indicator multiple cause (MIMIC)
- Mini-Mental State Examination (MMSE) 173–4
- minimal clinically important difference 441–3
- misfit indexes 169, 172
- missing at random (MAR) 364
- missing completely at random (MCAR) 363, 368–9
- missing data 66–7, 355–84
  - available-case analysis 356–7
  - biases 355, 360–2, 381, 383
  - complete-case analysis 356
  - compliance 383
  - degrees of freedom 381–2
  - EM algorithm 380
  - hierarchical scales 366
  - horizontal mean imputation 372–3

- missing data (*Continued*)
- hot deck imputation 379–80, 383
  - imputed data 355–8, 364–8, 369–81, 383
  - informative censoring 368
  - intermittent 358, 373
  - last value carried forward 370–1, 377–9
  - literature 384
  - longitudinal data 310–13, 317, 320, 352
  - Markov chain imputation 375–9, 383
  - missing forms 71, 358, 368–81, 387–8, 400
  - missing items 69–70, 362–4, 387, 400
  - multiple imputation 380–1
  - regression imputation 366–7, 373–5
  - reporting issues 387–8, 400
  - sample-mean imputation 371–2
  - sensitivity analysis 382
  - simple mean imputation 365–6
  - summary measures 357
  - types 358–9
- missing forms 71, 358, 368–81, 387–8, 400
- missing items 69–70, 362–4, 387, 400
- missing not at random (MNAR) 364, 369
- mixed-effects models 346
- MLEs *see* maximum likelihood estimates (MLEs)
- MNAR *see* missing not at random (MNAR)
- MOS *see* Medical Outcomes Study (MOS)
- MOS-HIV 84–5, 86–7, 104
- MPQ *see* McGill Pain Questionnaire (MPQ)
- MS *see* multiple sclerosis (MS)
- MTMM *see* multitrait–multimethod (MTMM) analysis
- multi-attribute utility measures 409–10
- multi-item scales 109–30
- causal items 127–9
  - causal variables 110
  - confidence intervals 113–14
  - construct validity 109, 116–23
  - correlations 111–16
  - Cronbach's  $\alpha$  123–9
  - dimensionality 109, 122–3
  - factor analysis 122–3
  - internal reliability 109, 123–8
  - overlap correction 116, 117
  - principles 31–2, 36–40
  - questionnaire design 62–3, 73
  - range of variables 112
  - rank correlations 115
  - scaling errors 118–19
  - significance tests 110–11, 112–13, 114–15
  - validation 109
- Multidimensional Fatigue Inventory (MFI-20) 11–12, 29, 68, 505–6
- multilevel models
- longitudinal data 329, 335, 347
  - reporting issues 390
- multiple
- endpoints 267
  - imputation 380–1
- multiple-indicator multiple cause (MIMIC) 43
- multiple sclerosis (MS) 103, 105
- multiplicity of outcomes 389–400
- multitrait-scaling analysis 117–22
- multitrait–multimethod (MTMM) analysis 87–91, 117
- Multivariate analysis of variance (MANOVA) 351–2, 391
- myocardial infarction 12
- necessary causes 45–6
- needs model 7
- NHP *see* Nottingham Health Profile (NHP)
- NNT *see* number needed to treat (NNT)
- nominal data 273–4
- non-compliance *see* compliance
- non-Normal distributions 275, 282–4, 288–90, 295
- norm-based scoring 19
- Normal distributions
- cross-sectional analysis 275, 280–1, 285–8, 295
  - factor analysis 150–3, 158
  - instrument selection 220
  - reporting issues 388
  - statistical tables 509–11

- norming systems 218–20
- norms *see* populations, norms
- Nottingham Health Profile (NHP)  
6, 20–1, 481  
compliance 228–9  
responsiveness 107
- null hypothesis 110–11
- number needed to treat (NNT) 452–4
- numerical data 275
- odds ratios (OR)  
clinical interpretation 431  
differential item functioning 180–1, 186  
meta-analysis 467–8, 470, 471  
reporting issues 388  
sample sizes 256, 258–60
- OHS *see* Oxford Hip Score (OHS)
- one-parameter logistic models 166, 170
- OR *see* odds ratios (OR)
- ordered categorical data  
cross-sectional analysis 274–5  
questionnaire design 58  
reliability 93–5  
sample sizes 258–60  
two-group comparisons 258–60  
*see also* Likert summated scales
- overlap  
correction 116, 117  
latent variables 83
- Oxford Hip Score (OHS) 221
- p*-values  
cross-sectional analysis 273, 281, 288  
reporting issues 390  
sample sizes 248
- PACIS 5
- Paediatric Asthma Quality of Life Questionnaire (PAQLQ) 26–7, 98–9, 501–2
- pain  
meta-analysis 466–8, 470  
sensitivity 103–5  
*see also* Brief Pain Inventory (BPI); McGill Pain Questionnaire (MPQ)
- Pain Anxiety Symptoms Scale (PASS) 274
- paired data  
longitudinal 341–2  
means 254–5, 262, 264  
proportions 256–7, 262, 265  
repeated measures 341–2
- palliative  
care 173–4, 346  
intent 9–11, 17
- PAQLQ 26–7, 98–9, 501–2
- parallel tests  
measurement scales 41–2, 47  
questionnaire design 53
- partial correlation 86
- partial-gamma 179, 181
- PASS *see* Pain Anxiety Symptoms Scale (PASS)
- path diagrams 134–5
- patient  
data 400  
information leaflets 241–2  
interviews 55  
preferences 8, 13–14, 17–19  
profiles, cross-sectional analysis 301–2  
variability 451–2
- Patient Generated Index (PGI) 7, 22–3, 487
- patient-orientated methods 427, 446–54
- Patient-Reported Outcome and Quality of Life Instruments Database (PROQOLID) 211
- patient-reported outcomes (PROs) 3–4, 211
- Pearson chi-squared test *see* chi-squared tests
- Pearson's correlation coefficient 96
- Pearson's *r* 111–12, 115
- perceived adjustment to chronic illness scale (PACIS) 5
- percentages  
clinical interpretation 430–1  
longitudinal data 308–9
- person-misfit indexes 169, 172
- PGI *see* Patient Generated Index (PGI)
- polychoric correlations 151
- polytomous item response models 171, 175–6

- POM *see* Profile of Mood States (POM)
- populations
- instruments 213–14
  - invariance requirement 195
  - norms 262–3, 427, 431–41
  - questionnaire design 52–3, 65–6
  - reporting issues 398
- post-treatment assessment 238, 243
- power 248
- pre-testing questionnaires 64–7
- precision 38–9
- predictive
- scales 46
  - validity 82–3
- preferences 406–9
- profile
- plots 301–2
  - scores 37
- Profile of Mood States (POM) 88–9
- proportions
- agreement 92
  - binomial 276–7
  - two-group comparisons 255–7, 262, 265
- PROQOLID *see* Patient-Reported Outcome and Quality of Life Instruments Database (PROQOLID)
- PROs *see* patient-reported outcomes (PROs)
- prostate cancer 375–80
- protocols
- administering QoL assessment 235
  - assessment during therapy 236–8
  - checklist 244–5
  - confidentiality 240–1
  - contact persons 234–5
  - follow-up assessment 238, 243
  - guidelines 235
  - help and proxy assessment 239–40
  - patient information leaflets 241–2
  - QoL assessment 233–4
  - randomisation 236, 242–3
  - randomised clinical trials 223, 225, 233–45
  - time frame 238–9
- proxy assessments 3–4, 5, 239–40
- psychometrics 40–5, 47
- psychosocial adaptation 14–15
- publication bias 474–5
- QALYs see* quality-adjusted life years (*QALYs*)
- QLQ-C30 23–4, 489–92
- clinical interpretation 432–4, 443
  - cross-sectional analysis 298, 302
  - differential item functioning 178, 183–5
  - factor analysis 141
  - graphical methods 298, 302
  - instrument selection 210, 212, 216–17
  - item response theory 203
  - longitudinal data 331–2, 351
  - measurement scales 32, 39
  - missing data 361–2, 365–6, 369–70
  - multi-item scales 116, 120–2, 125–6
  - questionnaire design 51, 54, 58, 61–2, 75
  - sample sizes 249
  - validation 89–91
- QOLIE-89 26, 497–500
- QOLQ for MS 103, 105
- Q-TWiST see* quality-adjusted time without symptoms and toxicity (*Q-TWiST*)
- qualitative methods 52, 55–8
- quality, meta-analysis 459–62
- Quality of Life in Epilepsy (QOLIE-89) 26, 497–500
- quality-adjusted life years (*QALYs*) 411–14
- quality-adjusted survival 405–25
- healthy-years equivalent 424–5
  - multi-attribute utility measures 409–10
- preferences 406–9
- Q-TWiST* 405, 414–25
- quality-adjusted life years 411–14
- sensitivity analysis 418–21
- utilities 406–14
- quality-adjusted time without symptoms and toxicity (Q-TWiST)* 405, 414–25
- questionnaire design
- field-testing 69–74
  - item generation 53–5
  - literature 75–6

- literature searches 53–4
- measurement scales 58–63
- pre-testing 64–7
- qualitative methods 52, 55–8
- target populations 52–3, 65–6
- translation 68–9
- validation 51–2, 64, 67–8
- wording of questions 63–4, 66, 70
- see also* instruments
  
- radiotherapy 15, 308–9
- random-effects models
  - meta-analysis 468–9, 472–3
  - repeated measures 344–6
- randomisation 236, 242–3
- randomised clinical trials (RCTs) 3, 223–45
  - administering QoL assessment 231–3, 235
  - baseline assessment 236
  - checklist 244–5
  - compliance 223–4, 225–30, 234
  - confidentiality 240–1
  - design issues 224–5
  - equivalence trials 224
  - follow-up assessment 238, 243
  - medical team issues 232, 235
  - organisational issues 225
  - patient information leaflets 241–2
  - patient issues 231–2, 235
  - protocols 223, 225, 233–45
  - proxy assessment 239–40
  - randomisation 236, 242–3
  - sample sizes 225, 247–69
  - selection criteria 17
  - significance tests 110
  - standard operating procedures 243–4
  - study type 224
- rank correlations 115
- ranked data *see* ordered categorical data
- Rasch models 166, 170, 171
- Rasch scores 221
- raw mean difference (MD) 464, 469–70
- raw scores 365
- RCTs *see* randomised clinical trials (RCTs)
- RE *see* relative efficiency (RE)
- reference values *see* populations, norms
- regression
  - imputation 366–7, 373–5
  - reliability 99
  - see also* linear regression; logistic regression
- relative efficiency (RE) 102–3, 105
- relative risk (RR) 467
- relative validity (RV) 102
- reliability
  - binary data 92–3
  - equality requirement 195
  - equivalent-forms reliability 91, 101
  - inter-rater reliability 91, 99–100
  - intra-class correlation coefficient 96–7
  - item response theory 195
  - measurement scales 38, 46
  - multi-item scales 109, 123–8
  - ordered categorical data 93–5
  - Pearson's correlation coefficient 96
  - single-item scales 78, 91–101
  - summated ratings 78, 91–101
  - test–retest reliability 46, 91, 95, 97–9
- repeatability 78, 91
- repeated measures
  - analysis of variance 335–8
  - auto-regression 338–40, 350
  - between- and within-subject variation 342
  - covariates 348–50
  - fixed-effects models 344–6
  - generalised estimating equations 342–4
  - interactions 341
  - longitudinal data 329, 335–50
  - mixed-effects models 345–6
  - modelling 336–8
  - multilevel models 347
  - paired data 341–2
  - random-effects models 344–6
  - reporting issues 390
- reporting issues 385–401
  - analytical tools 388–91, 400
  - association of variables 394
  - clinical significance 391
  - data analysis 386–91, 399–400
  - graphical methods 391–4, 396–7
  - guidelines 397–401
  - instruments 386, 399
  - literature 401

- reporting issues (*Continued*)  
 longitudinal data 322–6  
 means comparisons 389  
 missing data 387–8, 400  
 modelling techniques 390–1  
 multiplicity of outcomes 389–400  
 repeated measurements 390  
 sample sizes 386, 398–9  
 summary statistics 388  
 two group comparisons 392–4
- respondent validation 64
- response shifts 438–41
- responsiveness 78, 101–3, 105–7
- rheumatism 254, 257
- rotation 138–40, 147–9
- Rotterdam Symptom Checklist (RSCL)  
 25–6, 495–6  
 differential item functioning 176–7  
 factor analysis 153, 156–7  
 missing data 363
- RR *see* relative risk (RR)
- RSCL *see* Rotterdam Symptom Checklist (RSCL)
- RV *see* relative validity (RV)
- St John's Wort 463, 470–1, 475
- sample selection 57–8
- sample sizes 58, 247–69  
 attrition 268  
 choice of method 265–7  
 effect size 248–52  
 equivalence studies 263–5  
 estimating 248–52, 268  
 factor analysis 144–5, 154  
 literature 269  
 multiple endpoints 267  
 ordered categorical data 258–60  
*p*-values 248  
 power 248  
 randomised clinical trials 225  
 reference populations 262–3  
 reporting issues 386, 398–9  
 significance tests 248  
 target difference 267–8  
 time-to-event data 260–1, 262–3, 269  
 two-group comparisons 252–63
- sample-mean imputation 371–2
- scales *see* measurement scales; multi-item scales; single-item scales
- scaling errors 118–9
- Schedule for Evaluation of Individual Quality of Life (SEIQoL) 7, 22–3
- schizophrenia 348–9
- scree plots 145–6
- SEIQoL *see* Schedule for Evaluation of Individual Quality of Life (SEIQoL)
- selecting questionnaires 209–15
- selection bias 460–1
- self-reported health (SRH) 3
- self-reported questionnaires 17–19
- semi-structured interviews 55–6
- SEMs *see* structural equation models (SEMs)
- sensitivity 78, 101–5  
 analysis 382, 418–21
- separate calibration 194
- SF-36 14, 21–2, 482–4  
 cross-sectional analysis 276–7, 288  
 instrument selection 219–20  
 item response theory 169–71  
 measurement scales 39  
 multi-item scales 118–20, 125  
 questionnaire design 51  
 responsiveness 107  
 sample sizes 255, 264  
 sensitivity 103
- SG *see* standard gamble (SG)
- Sickness Impact Profile (SIP) 6, 20, 479–80  
 reliability 100  
 responsiveness 107  
 validation 80, 88–9
- significance tests  
 clinical interpretation 428–9  
 multi-item scales 110–11, 112–13, 114–15  
 sample sizes 248
- simple mean imputation 365–6
- simulations 200–1
- single-item scales 38–40, 73, 77–108  
 construct validity 77–8, 83–91  
 content validity 77–80  
 criterion validity 77–8, 81–3  
 face validity 79–80  
 item convergence 79  
 relevance 79

- reliability 78, 91–101
- repeatability 78, 91
- responsiveness 78, 101–3, 105–7
- sensitivity 78, 101–5
- SIP *see* Sickness Impact Profile (SIP)
- SMD *see* standardised mean difference (SMD)
- SOPs *see* standard operating procedures (SOPs)
- Spearman auto-correlation coefficients 330, 331–2
- Spearman–Browne prophecy formula 124
- Spearman's *r*, multi-item scales 115
- specialist interviews 54
- spurious correlation 86
- SRH *see* self-reported health (SRH)
- SRM *see* standardised response mean (SRM)
- standard gamble (SG) 407
- standard operating procedures (SOPs) 243–4
- standard scoring method 19
- standardised mean difference (SMD) 464–7, 471
- standardised response mean (SRM) 102, 107, 447–9, 466
- standardised scores 365
- statistical tables 509–15
- stopping rules 200–1
- structural equation models (SEMs) 43, 132, 158–60
- structured interviews 55–6
- Student's *t*-test 102–3, 252–4, 512
- sufficient causes 45–6, 218
- summary
  - measures 357
  - profiles 315–18
  - statistics 388
- summated ratings
  - construct validity 77–8, 83–91
  - content validity 77–80
  - criterion validity 77–8, 81–3
  - face validity 79–80
  - instrument selection 216–17
  - item convergence 79
  - questionnaire design 63
  - relevance 79
- reliability 78, 91–101
- repeatability 78, 91
- responsiveness 78, 101–3, 105–7
- sensitivity 78, 101–5
- validation 77–91
  - see also* Likert summated scales
- survival *see* Kaplan–Meier survival curves; quality-adjusted survival
- symmetry requirement 195
- T*-scores 218–20
- tabular presentation 320–2
- target populations *see* populations
- tau-equivalent tests 42
- TB *see* tuberculosis (TB)
- test
  - design 173–4
  - difficulty 162
  - equating/linking 191, 195–6
  - homogeneity 279
  - information functions 173, 191, 196–8
  - size 248
  - thought tests 35
- test-for-trend 279, 283, 286
- test–retest reliability 46, 91, 95, 97–9
- thought tests 35
- three-parameter logistic models 166, 171
- threshold
  - parameters 175–6
  - utility analysis 419, 421
  - values 430–1
- time trade-off (TTO) 407
- time-to-event data
  - cross-sectional analysis 285
  - sample sizes 260–1, 262–3, 269
- time without symptoms and toxicity (TWiST)* 405, 414–25
- transition
  - probabilities 375–8
  - questions 441
- translations
  - computer-adaptive testing 202
  - differential item functioning 183, 186
  - instruments 214
  - questionnaire design 68–9
- treatment-specific instruments 210–11

- TTO *see* time trade-off (TTO)
- tuberculosis (TB) 45–6
- TWiST see time without symptoms and toxicity (TWiST)*
- two-group comparisons  
  binomial proportions 276–7  
  categorical data 277–9  
  cross-sectional analysis 276–85  
  means 252–5, 262, 264  
  ordered categorical data 258–60  
  proportions 255–7, 262, 265  
  quality-adjusted survival 417–18  
  reference populations 262–3  
  reporting issues 392–4  
  sample sizes 252–63  
  time-to-event data 260–1, 262–3, 269
- two-parameter logistic models 166
- two-sample *t*-tests 252–4
- unidimensionality  
  constructs 33  
  item response theory 167–8  
  multi-item scales 109, 122–3
- unpaired  
  means 252–4, 262, 264  
  proportions 255–6, 262, 265
- utilities  
  instruments 410–11  
  measures 8  
  quality-adjusted survival 406–14
- validation 39–40, 77–91  
  construct validity 77–8, 83–91  
  content validity 77–80  
  criterion validity 77–8, 81–3
- instruments 213, 462  
  meta-analysis 462  
  multi-item scales 109  
  questionnaire design 51–2, 64, 67–8  
  single-item scales 77–91  
  summated ratings 77–91
- variability  
  longitudinal data 319  
  patient 451–2
- varimax rotation 139–40, 147–9
- VAS *see* visual analogue scales (VAS)
- visual analogue scales (VAS) 6, 39, 406  
  meta-analysis 464  
  questionnaire design 60–1
- Washington Psychosocial Seizure Inventory (WPSI) 80
- weighted  
  kappa coefficient 93–5  
  mean difference 469–70  
  sum scores 217–18
- WHO *see* World Health Organization (WHO)
- Wilcoxon test 389
- willingness to pay (WTP) 407
- within-person methods 443
- WMD *see* weighted, mean difference
- wording of questions 63–4, 66, 70
- World Health Organization (WHO) 4
- WPSI *see* Washington Psychosocial Seizure Inventory (WPSI)
- WTP *see* willingness to pay (WTP)
- Z-scores 218–20
- z-statistic 276, 289–90, 388



