

z Tests
for Proportions

	0.10	0.21	0.32	0.43	0.54	0.65	0.76	0.87	0.98	0.99
0.1	1.295556	0.959333	0.854355	0.779433	0.720400	0.675567	0.642567	0.618333	0.601200	0.594556
0.2	0.854355	0.675567	0.549433	0.464500	0.405467	0.360633	0.327633	0.303400	0.286267	0.279622
0.3	0.549433	0.405467	0.279433	0.194500	0.135467	0.090633	0.057633	0.033400	0.016267	0.009622
0.4	0.360633	0.279433	0.194500	0.135467	0.090633	0.057633	0.033400	0.016267	0.009622	0.002978
0.5	0.279433	0.194500	0.135467	0.090633	0.057633	0.033400	0.016267	0.009622	0.002978	0.000333
0.6	0.194500	0.135467	0.090633	0.057633	0.033400	0.016267	0.009622	0.002978	0.000333	0.000000
0.7	0.135467	0.090633	0.057633	0.033400	0.016267	0.009622	0.002978	0.000333	0.000000	0.000000
0.8	0.090633	0.057633	0.033400	0.016267	0.009622	0.002978	0.000333	0.000000	0.000000	0.000000
0.9	0.057633	0.033400	0.016267	0.009622	0.002978	0.000333	0.000000	0.000000	0.000000	0.000000
1.0	0.033400	0.016267	0.009622	0.002978	0.000333	0.000000	0.000000	0.000000	0.000000	0.000000
1.1	0.016267	0.009622	0.002978	0.000333	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1.2	0.009622	0.002978	0.000333	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1.3	0.002978	0.000333	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1.4	0.000333	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1.5	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1.6	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1.7	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1.8	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1.9	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2.1	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2.2	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2.3	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2.4	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2.5	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2.6	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2.7	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2.8	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2.9	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3.1	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3.2	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3.3	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3.4	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3.5	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3.6	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3.7	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3.8	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3.9	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
4.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

INDEX

- absolute effect measures, 61
- actuarial tables, 12, 15
- adaptive immune system, 110
- Aedes aegypti* mosquito vector (yellow fever), 19
- Agent Orange, 86
- age-specific display, 3, 177, 179
- age-specific rate standardization, 188, 190
- age-specific survival probabilities, 50
- Alzheimer's disease, 248–249
- amphetamine use, 171
- analyzing simple epidemiological data, 164–173
 - confidence intervals for measures of disease frequency, 164–166
 - confidence limits for incidence rate (example), 166
 - incidence rate data, 165–166
 - risk data (and prevalence data), 164–165
 - risk/prevalence, confidence limits (example), 165
- confidence intervals for measures of effect, 166–172
- case-control studies, 171
- cohort studies with incidence rate data, 169
- cohort studies with risk (prevalence) data, 167
- confidence limits for incidence rate difference, incidence rate ratio, 169–170
- confidence limits for risk difference, risk ratio, 168–169
- confidence limits for the odds ratio, 171–172
- P-value calculations, 172–173
- Andvord, Kristian, 20
- anecdotal information, limitations of, 4–5
- anesthetic gases (work of Snow), 15
- assumption of uniformity of effect, 189
- Atomic Bomb Casualty Commission, 80
- attack rate
 - cholera epidemic (example), 53, 70
 - defined, 42
 - secondary attack rate, 117
- attributable fraction measure, 27–28, 65–67
- Avicenna (Ibn Sina), 11
- basic reproductive number (R_0), 114–116
- bias, 124–145. *See also* confounding
 - defined, 124
 - induction time hypothesis and, 83
 - maternal recall bias, 133
 - random assignment and, 75
 - random vs. systemic errors, 124–126
 - recall bias, 19, 133–134
 - risk ratio and, 95
 - sources of
 - confounding, 33, 136–145
 - information bias, 133–136
 - selection bias, 105, 126–133
- biologic interaction
 - assessment of, with preventive factors, 208
- causal mechanism and, 202–205
 - defined, 202, 202–205
 - multiplicative relation as evidence, 201

INDEX

- absolute effect measures, 61
- actuarial tables, 12, 15
- adaptive immune system, 110
- Aedes aegypti* mosquito vector (yellow fever), 19
- Agent Orange, 86
- age-specific display, 3, 177, 179
- age-specific rate standardization, 188, 190
- age-specific survival probabilities, 50
- Alzheimer's disease, 248–249
- amphetamine use, 171
- analyzing simple epidemiological data, 164–173
 - confidence intervals for measures of disease frequency, 164–166
 - confidence limits for incidence rate (example), 166
 - incidence rate data, 165–166
 - risk data (and prevalence data), 164–165
 - risk/prevalence, confidence limits (example), 165
- confidence intervals for measures of effect, 166–172
- case-control studies, 171
- cohort studies with incidence rate data, 169
- cohort studies with risk (prevalence) data, 167
- confidence limits for incidence rate difference, incidence rate ratio, 169–170
- confidence limits for risk difference, risk ratio, 168–169
- confidence limits for the odds ratio, 171–172
- P-value calculations, 172–173
- Andvord, Kristian, 20
- anecdotal information, limitations of, 4–5
- anesthetic gases (work of Snow), 15
- assumption of uniformity of effect, 189
- Atomic Bomb Casualty Commission, 80
- attack rate
 - cholera epidemic (example), 53, 70
 - defined, 42
 - secondary attack rate, 117
- attributable fraction measure, 27–28, 65–67
- Avicenna (Ibn Sina), 11
- basic reproductive number (R_0), 114–116
- bias, 124–145. *See also* confounding
 - defined, 124
 - induction time hypothesis and, 83
 - maternal recall bias, 133
 - random assignment and, 75
 - random vs. systemic errors, 124–126
 - recall bias, 19, 133–134
 - risk ratio and, 95
 - sources of
 - confounding, 33, 136–145
 - information bias, 133–136
 - selection bias, 105, 126–133
- biologic interaction
 - assessment of, with preventive factors, 208
- causal mechanism and, 202–205
 - defined, 202, 202–205
 - multiplicative relation as evidence, 201

- biologic interaction (*Cont.*)
 - risk ratio and, 205–209
 - statistical interaction vs., 202
 - biomarkers, 29
 - birth defects, 56, 77, 153, 187–188
 - Black Death (bubonic plague) pandemic, 9, 112
 - blindness, 169–170
 - blinding, and use of placebos, 243–244
 - body mass index (BMI), 223
 - Boston Collaborative Drug Surveillance Program, 251
 - Bradford-Hill, Austin, 20
 - breast cancer 62, 80, 93–94, 168
 - Budd, William, 19, 20
 - Canon of Medicine* (Avicenna), 11
 - case-cohort sampling, 90
 - case-cohort studies, 96–98
 - data illustration, 98
 - random sampling of controls, 99
 - case-control studies, 87–109
 - blood type/breast cancer (example), 101
 - case-cohort studies, 96–98
 - case-crossover studies, 102–104
 - cohort study vs., 88, 106–107, 130–131
 - cross-sectional vs. longitudinal, 104–105
 - cumulative studies, 95–96
 - density studies, 90–94
 - hospital-based, 101
 - nested studies, 88–90
 - population-based, 99
 - prospective/retrospective studies, 101–102
 - referral population-based, 100
 - representation, importance of, 102
 - response rates, 105–106
 - retrospective cohort studies vs., 102
 - sources for controls, 87–88, 99–101
 - types of control sampling, 90
 - case-control study invention (Lane-Clayton), 18–19
 - case-crossover studies, 102–104
 - case fatality rate, 42, 114, 116, 167, 241
 - Cassel, John, 20
 - catalyst as cause, 29
 - causal criteria, 32–34
 - causal effect measures, 57–67
 - absolute vs. relative, 61
 - the counterfactual ideal, 57–58
 - effects measures, 58–61
 - attributable fraction, 65–67
 - examples, 61–63
 - risk ratios/rate ratios relations, 64–65
 - number of digits in report determination, 63
 - causal mechanism, for a given disease (“sufficient cause”), 24
 - biologic interaction and, 202–205
 - chance and, 149
 - complicated nature of, 198
 - component causes involved in, 24, 28
 - induction times, 28, 29, 81
 - interaction between causes, 26–27
- causal pie model (sufficient-component cause model), 24–35
- catalyst as cause, 29
- genetic vs. environmental causes, 25
- implications of, 24–29
 - induction time, 28–29
 - interaction between causes, 26–27
 - multicausality, 24–25
 - strength of causes, 25–26
 - sum of attributable fractions, 27–28
- scientific inference process, 30–35
 - causal criteria, 32–34
 - generalization in epidemiology, 34–35
 - induction, 30–31, 31–32
- causation. *See also* causal effects measures; multicausality
 - approach of assigning, 57
 - Avicenna's contributions, 11
 - described, 23–24
 - Hippocrates' contributions, 10
 - significance testing and, 161
- Chagas disease, 113
- characteristics, prevalence of, 56
- chicken and egg, rate problem example, 46
- cholera, 62
- cholera epidemic
 - attack rate, 53, 70
 - Farr's analysis of, 15
 - oral-fecal transmission, 114
 - population at risk study and, 74–75
 - Snow's analysis of, 15–16, 19, 52, 70–71

- clinical settings for use of epidemiology, 235–252
 - comparative effectiveness, 235
 - diagnosis
 - “gold standard” disease definition, 235–236
 - length-biased sampling, 240–241
 - predictive value, 237–239
 - screening, 239–240
 - sensitivity and specificity, 236–237
- health-outcomes research, 235, 252
- prognosis, 241–242
- therapy
 - clinical trials, 72–73, 242–249
 - pharmacoepidemiology, 228, 235, 249–251
- clinical trials
 - blinding and use of placebos, 243–244
 - described, 72–73
 - ethics of placebo use, 244
 - field trials vs., 73–74, 121
 - null hypothesis, unrejectability, 248–249
 - random assignment in, 72
 - threats to validity
 - confounding imbalances, 246–247
 - incomplete follow-up, 245
 - intent-to-treat analysis, 245–246
- closed cohorts, 77–78, 79, 81
- clozapine, 184–186, 191
- cohort studies, 69–87
 - case-control study vs., 88, 106–107, 130–131
 - closed cohorts, 77–78, 79, 81
 - counting disease events, 78–79
 - defined/described, 69–70
 - determining measures to report, 84
 - eligibility criteria, 83–85
 - example
 - atomic bomb (Japan) survivors, 80
 - vitamin A during pregnancy, 76–77, 87
 - x-ray fluoroscopy, breast cancer, 80
 - experiments, 71–74
 - exposure and induction time, 80–83
 - exposure classification, 83–85
 - incident rates/risks measurements, 79–80
 - open cohorts, 77–78
 - population at risk in, 74–76
 - retrospective cohort studies vs., 85–86, 102
 - special exposure, general population, 86–87
 - time loops, 83–85
 - tracing of subjects, 86
- Collegium Ramazzini (Carpi, Italy), 14
- common sense vs. epidemiologic point of view, 4
- communicable (defined), 117
- community concerns
 - drinking water treatment, 8–9
 - transmissible disease control, 9
- community intervention trial, 74
- comparative effectiveness, 235
- competing risk, 42, 50, 79, 80, 204
- component causes, 24
- conditional logistic model, 195
- confidence interval function, 152–157
- confidence intervals
 - calculation of, 160–162
 - confidence limits vs., 169
 - defined, 149, 150
 - estimation using, 161
 - formation of, 156
 - functions of, 152–157
 - P-value function relation to, 156
 - relation to point estimates, 149–151
- confidence intervals for measures of
 - disease frequency, 164–166
 - confidence limits for incidence rate (example), 166
 - incidence rate data, 165–166
 - risk data (and prevalence data), 164–165
 - risk/prevalence, confidence limits (example), 165
- confidence intervals for measures of effect, 166–172
 - case-control studies, 171
 - cohort studies with incidence rate data, 169
 - cohort studies with risk (prevalence) data, 167
 - confidence limits for incidence rate difference, incidence rate ratio, 169–170
 - confidence limits for risk difference, risk ratio, 168–169

- confidence intervals for measures of effect (*Cont.*)
 - confidence limits for the odds ratio, 171–172
- confidence limits, 156, 160
 - for an incidence rate, 166
 - confidence interval vs., 169
 - for incidence rate difference, incidence risk ratio, 169–170
 - for the odds ratio, 171–172
 - for risk difference, risk ratio, 168–169
- confounding, 4, 33, 130, 136–145, 193–194. *See also* stratification of data, for controlling confounding
 - control of, with regression models, 145, 218–221
 - defined, 136
 - diabetes treatment program example, 140–142
 - Down syndrome example, 136–139, 176
 - factor properties, 141
 - by indication, 137, 248–249
 - measurement of, 193–194
 - methods of control, 142–145
 - matching, 144–145
 - randomization, 142
 - restriction, 142, 142–144
 - mortality data example, 2–3
 - overfitting regression models with
 - summary confounder scores, 226–228
 - problematic aspect of, 4
 - in a randomized experiment, 143
 - residual confounding, 181, 227, 231
 - trigeminal neuralgia mortality rate example, 176–177
 - variable matching ratios, trimming and, 229–230
- conjecture, 31, 34
- contact tracing studies, 120
- De contagione et contagiosis morbis et curatione* (Fracastoro), 11
- control sampling, in population-based studies, 99
- control selection, in density-case control studies, 92–93
- counterfactual study, 57–58
- cross-sectional studies
 - described, 104
 - longitudinal studies vs., 104–105
 - prevalence data and, 167
- crossover studies, 57
 - case-crossover studies, 102–104
- crude data
 - in case-control studies, 187
 - confounding and, 140, 141, 164, 176, 194
 - defined, 3
 - matching and, 130
 - stratification of risk data and, 179
- cumulative case-control studies, 95–96
- cumulative sampling, 90, 96
- data collection
 - Farr's work, 14–15
 - Graunt's work, 12–13
- death
 - Bills of Mortality classification, 13
 - case fatality rate, and attack rate, 42
 - mortality rate measure, 43–44
 - polar-area diagram plotting of, 17–18
 - risks of, 5–6
 - of subjects in studies, 40–41
- Declaration of Helsinki, 244
- dengue fever, 113
- density-based sampling, 90, 93
- density case-control studies, 90–94
 - case-control data illustration, 93–94
 - control selection in, 92–93
 - odds ratio, 95
- depression, 158–160
- diabetes, 140–141, 179–180
- diethylstilbestrol, 28, 40, 87
- direct standardization, 192
- disease events, 78–79, 82, 86
- disease occurrence measures
 - chicken and egg rate problem example, 46
 - incidence rate, 42–47
 - point-source, propagated epidemics, 51–53
 - prevalence proportion, 53–56
 - risk and incidence proportion, 38–42
 - risk and incidence rate relation, 47–51
- Doll, Richard, 20
- domestic violence, 40–41
- double-blind trials, 243

- Down syndrome example of confounding, 136–139, 176
- drug-eluting stents, 228–232
- dynamic cohort. *See* open cohort
- dynamic population. *See* open cohort
- Ebola virus, 114, 241
- effect-measure modification, 199–201
 - arbitrariness and, 200, 202
 - effect modification vs., 200
- effective reproductive number (R_e), 116, 120
- Eisenberger, M. A., 157–158
- eligibility criteria, cohort studies, 83–85
- endemic equilibrium, 116
- environmental causes, 25
- epidemic. *See also* cholera epidemic
 - curve shape factors, 52–53
 - defined, 51
 - point-source epidemic, 51–53, 114
 - propagated epidemic, 53
- epidemic hysteria, 53
- epidemiologic studies
 - case-control studies, 87–109
 - case-cohort studies, 96–98
 - case-crossover studies, 102–104
 - cohort compared with, 106–107
 - cross-sectional vs. longitudinal, 104–105
 - cumulative studies, 95–96
 - density studies, 90–94
 - nested, 88–90
 - prospective/retrospective studies, 101–102
 - representation, importance of, 102
 - response rates, 105–106
 - sources for controls, 99–101
 - types of control sampling, 90
 - cohort studies, 69–87
 - case-control with, 106–107
 - closed and open cohorts, 77–78
 - counting disease events, 78–79
 - defined/described, 69–70
 - eligibility criteria, 83–85
 - experiments, 71–74
 - exposure and induction time, 80–83
 - exposure classification, 83–85
 - incident rates/risks measurements, 79–80
- epidemiologic thinking
 - age-specific display/crude data, 3
 - anecdotal information, 4–5
 - confounding, 2–3, 4
 - population pyramid, 1–2
 - risk, 5–6
 - stratification, 3
- equine encephalitis (vector-borne zoonoses), 114
- ergonomics, Ramazzini's interest in, 14
- estimation, 148–149
 - statistical hypothesis testing vs., 151–152
 - statistical significance testing vs., 161
 - using confidence intervals, 161
- expectation of life, 46
- experiments (cohort study), 71–74
 - as imperfect gold standard, 75
 - natural experiments vs., 73–74
 - types of
 - clinical trials, 72–73
 - community intervention trial, 74
 - field trials, 73–74
- exposed cohorts, 69
- exposure and induction time (in cohort studies), 80–83
- exposure classification, in cohort studies, 83–85
- factoring of exposure, 225
- Farr, William, 14–15
- Feyerabend, P., 32
- field trials
 - described, 73–74
 - Salk polio vaccine (example), 74
- fluoridation, 74
- flutamide, 157–158
- natural experiment (Snow), 70–71
- population at risk, 74–76
- retrospective cohort studies, 85–86
- special exposure, general population, 86–87
- time loops, 83–85
- tracing of subjects, 86
- vitamin A during pregnancy (example), 76–77
- x-ray fluoroscopy, breast cancer (example), 80
- reasons for use of models, 211

- Framingham Heart Study, 78
 Francastoro, 11–12
 Frost, Wade Hampton, 19–20, 118–119
 general linear regression model, 213–215
 transformation of, 213–215
 variables, dependent/independent, 213, 213–214
 general-population cohort studies, 86–87
 generalization, in epidemiology, 34–35
 generation time (defined), 117, 118
 genetic causes, 25
 germ theory
 Fracastoro's early work, 11–12
 miasmas and, 16
Girardia zoonotic protozoan parasite, 114
 Goldberger, Edward, 20
 Graunt, John, 12–13
 Greenwood, Major, 20
 Haack, S., 32
 health-outcomes research, 235, 252
 healthy worker effect, 127
 herd immunity, 20, 114–116
 heterogeneity test, 182
 Hill, A. B. ("Hill criteria"), 32–34
 Hippocrates, 10, 11
 historical background of epidemiology, 9–20
 Avicenna (Ibn Sina), 11
 Farr, William, 14–15
 Francastoro, 11–12
 Frost, Wade Hampton, 19–20
 Graunt, John, 12–13
 Hippocrates, 10
 Nightingale, Florence, 17–19
 Ramazzini, Bernardino, 14
 Simmelweis, Ignasz, 16–17
 Snow, John, 15–16
 historical cohort studies. *See* retrospective cohort studies
A History of Public Health (Rosen), 8
 HIV infection, 73
 H1N1 influenza pandemic (1918–19), 19–20, 111–112, 112
 homogeneity test, 182
 hospital-based case-control studies, 101
 host population, as *reservoir* for pathogens, 112
 Hume, David, 30–31
 hypertension, 206–207
 hysteria outbreak, 53–54
 immortal person-time, 84–85
 immunity to infectious agents, 111, 117
 incidence rate, 42–47
 as "annual incidence," 45–46
 calculation, in case-control study, 91
 defined, 42–43
 measurements in cohort studies, 79–80
 mortality rate (defined), 43–44
 odds ratio calculation, 94
 person-time time value, 45
 relation of risk and, 47–51
 exponential decay, 48
 life-tables, 50
 survival analysis, 49
 time at risk of disease component, 43
 waiting time measure, 46
 incubation period (defined), 117
 indirect standardization, 192
 induction method (of gaining insight), 30–31
 Hume's skepticism of, 30–31
 induction time, 28–29
 induction period, 28, 29
 induction time, 28–29
 infant mortality, 15, 16–17
 infectious disease epidemiology, 110–123
 basic reproductive number (R_0), 115–116
 effective reproductive number (R_e), 116, 120
 endemic equilibrium, 116
 herd immunity, 114–116
 host defenses of humans, 111
 immunity to infectious agents, 111
 outlook for, 121–122
 public health burden of disease, 111–112
 Reed-Frost epidemic model, 20, 53, 118–119
 types of investigations, 119–121
 contact tracing studies, 120
 outbreak investigations, 120
 seroprevalence surveys, 120–121
 vaccine trials, 121
 types of transmission, 112–114
 isolation, quarantine strategy, 116
 person-to-person, 113, 114–116

- point source, propagated, 114
 vector-borne, 113
 zoonoses, 114
 inference, canons of, 32
 influenza, 19. *See also* H1N1 influenza pandemic (1918–19)
 attack rate, 42
 case-control vaccine study, 90, 121
 causal interaction and, 198
 as a propagated epidemic example, 53, 114
 risk data example, 164
 selection bias in studies, 127
 zoonose method of spreading, 114
 information bias, 133–136
 differential misclassification, 133, 133–134
 nondifferential misclassification, 133, 134–135
 recall bias, 19, 133–134
 initiator (reference to component causes of cancer), 28
 innate immune system, 110
 instantaneous risk ratio, 65
 intent to treat analysis (in clinical trials), 72–73
 interactions. *See also* biologic interaction; statistical interaction
 causal interactions/public health, 198
 between causes, 26–27
 consequences of, 27
 defined, 199
 immunity-social, during disease outbreak, 116
 interactions, measurement of, 198–209
 effect-measure modification, 199–201
 arbitrariness and, 200, 202
 effect modification vs., 200
 partitioning risk in joint exposure, 205–209
 intermittent claudication, 221–222
 isolation and quarantine strategy, 116
 Jenner, Edward, 20
 Johns Hopkins Bloomberg School of Public Health, 19
 Kaplan-Meier product-limit method, 241
 Kuhn, Thomas, 32
 Lane-Claypon, Janet, 18–19
 laryngeal cancer, 211–213
 latent period, 29, 81
 length-biased sampling, 240–241
 leukemia, 81, 100–101, 174
 life table, 50–51
 Lilienfeld, Abraham, 20
 Lind, James, 20
 logistic transformation, 215–216
 longitudinal studies vs. cross-section studies, 104–105
 Louis, Pierre, 20
 Love Canal, 86
 lung cancer, 24, 25, 53, 65, 108, 198, 201, 205–206
 Lyme disease, 113
 Maclure, M., 103
 MacMahon, Brian, 20
 magnetic fields, 174
 Malthus, Thomas Robert, 15
 Mantel, N., 136
 Mantel-Haenszel formulas, 184, 189, 191
 matching
 in case-control studies, 131, 129–132
 vs. cohort studies, 230
 confounding and, 130, 144–145, 195
 convenience of, 100
 crude data and, 130
 drawbacks vs. advantages, 132
 motivation for, 129
 in propensity score models, 229
 restriction vs., 144–145
 risk ratio (RR) and, 129–130
 risk-set sampling and, 132
 selection bias and, 131
 stratification after, 195
 variable matching ratios, 229–230
 maternal recall bias, 133
 Mauri, L., 228–232
 measures of disease occurrence
 miasmas
 Farr's belief in, 16
 Simmelweis's belief in, 17
 Mill, John Stuart, 32
 Milunsky, A., 76
 models, reason for use in epidemiology, 211
 mortality rate
 age-standardized rate, 211



- mortality rate (*Cont.*)
 changes with time, over time scales, 46
 clozapine drug use study, 184
 defined, 43-44
 Frost's birth cohort analysis of, 20
 selection bias factor in studies, 127
 Snow's cholera "natural experiment" study, 70
- motor vehicle injury, 50-51
- multicausality, 24-25
- multivariable regression modeling, 211, 219, 221
- mutualism, 110
- mysticism and religion vs. observation and reason, 10
- National Institutes of Health, 27
- natural experiment, 16, 73
- neonatal mortality, 74
- nested case-control studies, 88-90
- Nightingale, Florence, 17-19
- null hypothesis
 clinical trials and, 248-249
 defined, 150-151
 and P-values, 150, 152, 160
 probability of correctness, 151
 significant test and, 155
- occupational epidemiology, 88
- occupational hazards
 Farr's work on, 15
 Ramazzini's work on, 14
- odds ratio, 84, 91, 95, 96, 98
- open cohort, 77-78
- oral contraceptives, 206-207
- outbreak investigations, 120
- outcomes research. *See* health-outcomes research
- overfitting regression models with summary confounder scores, 226-228
- P-value functions, 152-157
 data interpretation, 154-155
 examples
 drug exposure during pregnancy, congenital heart disease study, 152-154
 randomized trial of flutamide, 157-158
- St. John's Wort effectiveness, 158-160
- hypothetical case-control data, 154
- relation to confidence intervals for a given estimate, 156
- P-values
 calculation of, 150-151
 case control data, 173
 incidence rate data, 172-173
 risk data, 172
 for stratified data, 192-193
 null hypothesis and, 150, 152, 160
 statistical significance determination, 151-152
- pandemic (defined), 112. *See also* Black Death (bubonic plague)
 pandemic; influenza
- Panum, Peter, 20
- parasitism, 110
- person-to-person transmission of infection, 113, 114-116
- pharmacoepidemiology
 advances/described, 235
 post-marketing surveillance, 250
 spontaneous reports to the FDA, 250
 use of health databases, 251
 use of propensity scores, 228
- placebos
 blinding and use of, 243-244
 depression/St. John's Wort vs., 158
 diabetes trial/tolbutamide vs., 140, 179
 ethics and use of, 244
 HIV/zidovudine vs., 73
 prostate cancer/flutamide vs., 157
 ramipril (ACE inhibitor) vs., 242
 Salk vaccine trial, 41
- plagues. *See* Black Death (bubonic plague)
 pandemic; cholera epidemic
- Plasmodium* protozoa, 113, 122
- point estimate, 149
- point source epidemic, 51-53, 114
- polar-area diagram (coxcomb), of Nightingale, 17-18
- poliomyelitis, 19, 74, 121, 122
- pooling
 standardization vs., 189-191
 stratified analysis use of, 178-179, 189-191, 201
- Popper, Karl, 31-32. *See also* refutationism

- population at risk in cohort studies, 74-76
- population-based study, 99
- population growth theories (Malthus), 15
- population pyramid, 1-2
- post-marketing surveillance, 250
- predictive value, 237-239
- prevalence odds, 55
- prevalence proportion, 53-56
 application areas, 56
 defined, 53-54
 factors influencing, 54-55
 prevalence characteristics, 56
 prevalence odds, 55
 steady state factor, 55
- preventive medicine, 8
- primary cases (defined), 42
- prognosis, 241-242
 clinical epidemiology and, 235
 confounding by indication and, 243
 disease screening and, 239
 length-bias sampling and, 240
- promotor (reference to component causes of cancer), 28
- propagated epidemic, 53, 114
- propensity score models
 defined, 227
 matching in, 229
- pharmacoepidemiologic use, example, 228-232
 selecting variables for, 228
- prospective/retrospective case-control studies, 101-102
- prostate cancer, 157-158
- proxy sample, 90, 101
- public health
 Avicenna's influence on, 11
 Hippocrates' influence on, 10
 Nightingale's work, 17-19
 origins of, 8-9
- quarantine and isolation strategy, 116
- rabies zoonotic virus, 114
- Radiation Effects Research Foundation, 80
- Ramazzini, Bernardino, 14
- random assignment
 confounding and, 75, 248
 described, 72
 intent-to-treat approach vs., 72-73, 245
- internal validity and, 106
- possible study imbalances, 140
- in randomized trials, 148
- types of results produced by, 248
- usage in experiments, clinical trials, 142, 243
- random error. *See also* random error, and the role of statistics
 described, 143, 148
 systematic error vs., 124-126
- random error, and the role of statistics, 148-162. *See also* P-value functions
- chance, 149
- estimation, 148-149
 statistical hypothesis testing vs., 151-152
 statistical significance testing vs., 161
- null-hypothesis, 150-151
- probability of correctness, 151
- P-values, 150-151
 and statistical significance determination, 151-152
- point estimates, confidence intervals, 149-151
- role of statistics, 148
- statistical hypothesis testing, 152
- randomized trials, 59, 148
- rare disease assumption, 96
- recall bias, 19, 133-134
- Reed, Lowell, 20, 118-119
- Reed-Frost epidemic model, 20, 53, 118-119
- referral population-based, case-control studies, 100
- refutationism, 31-32
 conjecture component, 31
 criticism of, 32
- regression line (defined), 211
- regression models in epidemiologic analysis, 211-233
 advantages of, 226, 232
 centering of variables in, 223
 choices among models, 216-218
 control of confounding with, 145, 218-221
 factoring of exposure, 225
 general linear model, 213-215



regression models in epidemiologic analysis (*Cont.*)
 linear regression, 201
 logistic regression, 195, 201
 logistic transformation, 215–216
 matching in case-control studies and, 195
 multivariable models, 211, 219, 221
 overfitting with summary confounder scores, 226–228
 overreliance on, 232
 predicting risk for a person, 221–222
 propensity scores, 227
 selection bias and, 131
 simple regression, 212
 spline regression, 225
 stepwise models, 224–226
 strategy for construction of, 222–224
 stratification comparison, 196, 220, 232–233
 variable matching ratios, confounding, trimming, 229–230
 relative effect measures, 61
 relative risk (RR)
 in case-control studies, 150
 described, 65
 representativeness, 34–35, 102
 reproductive number (defined), 117
 reservoir for pathogens, 112, 117, 122
 residual confounding, 181, 227, 231
 response rates, in case-control studies, 105–106
 retrospective cohort studies, 85–86
 risk
 of breast cancer, 19
 of death, 5–6
 incidence rate relation to, 47–51
 exponential decay, 48
 life-tables, 50
 survival analysis, 49
 measures of, 38–39
 in cohort studies, 79–80
 competing risk issue, 40–41
 interpretation, 39–40
 loss to follow-up issue, 41
 occupational hazards (Ramazzini), 14
 predicting, for a person, 221–222
 risk not meaning risk, 65

risk ratios/rate ratios, relation between, 64–65
 risk-odds ratio, 215, 216, 218, 221
 risk ratio (RR), 60, 61
 attributable fraction calculation from, 66
 bias-caused overestimation, 95
 biologic interaction and, 205–209
 case-cohort study estimates, 97
 cumulative sampling and, 96
 instantaneous risk ratio, 65
 matching/confounding and, 129–130
 odds ratio relation to, 84, 96, 97
 rate ratio relation to, 64–65, 95, 96
 reporting of, 63
 St. John's Wort depression example, 159
 risk-set sampling, 93, 97, 132
 Rose, Geoffrey, 20
 Rosen, George, 8
 rounding, 63
 Salk polio vaccine, field trial (example), 41, 74
 SARS (Severe Acute Respiratory Syndrome), 116
 scientific inference process, 30–35
 causal criteria, 32–34
 generalization in epidemiology, 34–35
 induction, 30–31, 31–32
 screening, 34, 126, 239–241
 secondary attack rate (defined), 42, 117
 secondary cases (defined), 117
 selection bias, 105, 126–133
 defined, 126
 healthy worker effect, 127
 influenza vaccine efficacy example, 127–128
 matching in case-control studies, 129–132
 self-selection bias and, 126–127
 self-selection bias, 126–127
 Semmelweis, Ignasz, 16–17
 sensitivity, 236–239
 seroprevalence surveys, 120–121
 Shelton, R. C., 158
 simple regression (defined), 212
 smoking, 2–4, 24, 25, 26, 53, 83, 201, 205–206, 211–213

Snow, John, 15–16, 19, 52
 natural experiment of, 70–71, 74–75, 98
 special-exposure cohort studies, 86–87
 specificity, 236–239
 spermicides, 187–188
 spline regression, 225
 spontaneous reports (to the FDA), 250
 St. John's Wort (*Hypericum perforatum*),
 effectiveness in relieving depression (P-value function example), 158–160
 standard error, 151, 160, 165
 standardization, 188–192
 assumption of uniformity of effect, 189
 crude vs. standardized rates, 190
 defined, 188
 indirect/direct, 192
 male/female, age-specific weights (example), 188–189
 pooling vs., 189–191
 reasons for, 189
 standardized mortality (or morbidity) ratio (SMR), 192
 standardized rate difference, 191
 standardized rate ratio, 191
 Stark, C. R., 136
 statistical hypothesis testing, 152
 statistical hypothesis vs. estimation, 151–152
 statistical interaction, 199, 200, 202, 207
 statistical significance, 151–152, 161
 steady state, 55
 stents, 228–232
 stepwise models in epidemiologic analysis, 224–226
 exposure-disease relation shape estimation, 224–225
 interaction evaluation, 226
 stratification of data
 age-specific display (stratified by age), 3, 177, 179
 defined, 140
 importance of, 195–196
 regression analysis comparison, 196, 220
 use in cohort studies, 84
 stratification of data, for controlling confounding, 176–196
 after matching, 195

calculation of P-values, 192–193
 diabetes study (example), 140–141
 Down syndrome (example), 136–139, 176
 measuring confounding, 193–194
 standardization, 188–192
 stratification after matching, 195
 trigeminal neuralgia mortality rate example, 176–177
 two or more variables, 194–195
 unconfounded effect estimates,
 confidence intervals, 178–181
 case-control studies, 187–188
 cohort studies with incidence rate data, 184–186
 cohort studies with risk, prevalence data, 179
 confidence intervals, for pooled estimates, 182–184
 pooling, 178–179, 189–191
 stratification of risk data (example), 179–182
 stratum-specific estimates, 178
 stroke, 171, 206–207
 sufficient-component cause model.
 See causal pie model (sufficient-component cause model)
 sum of attributable fractions, 27–28
 Surgeon General's Report on Smoking and Health (1964), 32–34
 survival analysis, 49–51
 susceptibility, 117, 119, 202
 Sydenstricker, Edgar, 20
 symbiosis (defined), 110
 systematic error. See bias
 tamoxifen, 168–169
 time at risk of disease (in incident rate calculation), 43
 time loops, in cohort studies, 83–85
 toxic shock syndrome, 32, 120
 toxoplasmosis, 114
 tracing of subjects, in cohort studies, 86
 transmissible disease control, 9
 transmission probability, 117
 tuberculosis, 236–238
 triple-blind trials, 243
 unexposed cohorts, 69

profound to
 design, and
 cations that
 goal of pro-
 "epidemiology
 introductory
 vity through
 selectivity of
 es an excep-
 xt, graphics,
 eful in expli-
 design, and
 course . . . The
 r authors and
 ology Monitor
 ologist on the
 unify the uni-
 epidemiology



- vaccines, 121
 - trials, 41, 57, 121
 - influenza case-control study, 90, 121
 - Salk polio vaccine field trial, 74
 - selection bias in, 127–128
- variable matching ratios, 229–230
- variables
 - bias and study variables, 124
 - categorical variables, 145
 - centering of, in regression models, 223
 - confounding variables, 93, 136, 139, 142, 176, 178, 193
 - dependent/independent, 213
 - disease/exposure variables, 104
 - information bias and, 133
 - matching/matched variables, 131
 - nominal scale variables, 178
 - in selection bias, 106, 131
 - selection for propensity score models, 228
 - stratification variables, 178, 189, 194–195, 201
 - time variable, in risk-set sampling, 132
- vector-borne transmission of infections, 113, 117
- Vienna Maternity Hospital, 16–17
- virulence, 117
- vitamin A during pregnancy (cohort study), 76–77, 87
- waiting time, 46
- Weiss, N. S., 34
- West Nile encephalitis, 113
- Wilson's confidence limits, 174
- World Health Organization (WHO), 112
- x-ray fluoroscopy/breast cancer cohort study (example), 80
- zoonoses, 114