ISO/TS 20281:2006-04 (E)

Water quality - Guidance on statistical interpretation of ecotoxicity data

Contents

Page

Foreword xii

Introduction xiii

1 Scope 1

2 Normative references 1

3 Terms and definitions 1

4 General statistical principles 8
4.1 Different statistical approaches 8
4.1.1 General 8
4.1.2 Hypothesis-testing methods 8
4.1.3 Concentration-response modelling methods 10
4.1.4 Biology-based methods 11
4.2 Experimental design issues 11
4.2.1 General 11
4.2.2 NOEC or ECx: Implications for design 12
4.2.3 Randomization 12
4.2.4 Replication 13
4.2.5 Multiple controls included in the experimental design 13
4.3 Process of data analysis 14
4.3.1 General 14
4.3.2 Data inspection and outliers 14
4.3.3 Data inspection and assumptions 15
4.3.3.1 Scatter 15
4.3.3.2 Heterogeneous variances and distribution 15
4.3.3.3 Heterogeneous variances and true variation in response 16
4.3.3.4 Consequences for the analysis 16
4.3.4 Transformation of data 16
4.3.5 Parametric and non-parametric methods 17
4.3.5.1 General 17
4.3.5.2 Parametric methods 17
4.3.5.3 Generalized linear models (GLMs) 18
4.3.5.4 Non-parametric methods 18
4.3.5.5 How to choose? 18
4.3.6 Pre-treatment of data 19
4.3.7 Model fitting 19
4.3.8 Model checking 20
4.3.8.1 Analysis of residuals 20
4.3.8.2 Validation of fitted dose-response model 21
4.3.9 Reporting the results 21

5 Hypothesis testing 21
5.1 Introduction 21
5.1.1 General 21
5.1.2 NOEC: What it is, and what it is not 25
5.1.3 Hypothesis used to determine NOEC 25
5.1.3.1 Understanding the question to be answered 25
5.1.3.2 One-sided hypothesis 26
5.1.3.3 Two-sided trend test 26

5.1.3.4 Trend or pair-wise test 26
5.1.4 Comparisons of single-step (pair-wise comparisons) or step-down trend tests to determine the NOEC 28
5.1.4.1 General discussion 28
5.1.4.2 Single-step procedures 28
5.1.4.3 Step-down procedures 29
5.1.4.4 Deciding between the two approaches 30
5.1.5 Dose metric in trend tests 31
5.1.6 Role of power in toxicity experiments 31
5.1.7 Experimental design 32
5.1.8 Treatment of covariates and other adjustments to analysis 33
5.2 Quantal data (e.g. mortality, survival) 34
5.2.1 Hypothesis testing with quantal data to determine NOEC values 34
5.2.2 Parametric versus non-parametric tests 35
5.2.2.1 Basis 35
5.2.2.2 Single-step procedures 36
5.2.2.3 Step-down procedures 36
5.2.2.3.1 Choice of step-down procedure 36
5.2.2.3.2 Test for monotone dose response 36
5.2.2.3.3 Analysing the monotonic response for quantal data -- Step-down procedure 37
5.2.2.3.4 Possible modifications of the step-down procedure 37
5.2.2.4 Alternative procedures 37
5.2.2.4.1 Parametric and non-parametric procedures 37
5.2.2.4.2 Pair-wise ANOVA-based methods 38
5.2.2.4.3 Jonckheere-Terpstra trend test 38
5.2.2.4.4 Poisson tests 38
5.2.2.5 Assumptions of methods for determining NOEC values 38
5.2.3 Additional information 39
5.2.4 Statistical items to be included in the study report 40
5.3 Hypothesis testing with continuous data (e.g. mass, length, growth rate) to determine NOEC 40
5.3.1 General 40
5.3.2 Parametric versus non-parametric tests 41
5.3.3 Single-step (pair-wise) procedures 42
5.3.3.1 General 42
5.3.3.2 Dunnett's test 42
5.3.3.3 Tamhane-Dunnett test 42
5.3.3.4 Dunn's test 42
5.3.3.5 Mann-Whitney test 43
5.3.4 Step-down trend procedures 43
5.3.5 Determining the NOEC using a step-down procedure based on a trend test 43
5.3.5.1 General 43
5.3.5.2 Preliminaries 43
5.3.5.3 Step-down procedure 43
5.3.5.3.1 Preferred approach 43
5.3.5.3.2 Williams' test 44
5.3.5.3.3 Jonckheere-Terpstra test 44
5.3.6 Assumptions for methods for determining NOEC values 44
5.3.6.1 Small samples -- Massive ties 44
5.3.6.2 Normality 45
5.3.6.3 Variance homogeneity 45
5.3.7 Operational considerations for statistical analyses 46
5.3.7.1 Treatment of experimental units 46
5.3.7.2 Identification and meaning of outliers 46
5.3.7.3 Multiple controls 46
5.3.7.4 General 47
5.4 Statistical items to be included in the study report 47

6 Dose-response modelling 48
6.1 Introduction 48
6.2 Modelling quantal dose-response data (for a single exposure duration) 49
6.2.1 General 49

6.2.2 Choice of model 50
6.2.2.1 General 50
6.2.2.2 Probit model 51
6.2.2.3 Logit model 53
6.2.2.4 Weibull model 54
6.2.2.5 Multi-stage models 55
6.2.2.6 Definitions of EC50 and ECx 55
6.2.3 Model fitting and estimation of parameters 56
6.2.3.1 Software and assumptions 56
6.2.3.2 Response in controls 56
6.2.3.3 Analysis of data with various observed fractions at each dose group 57
6.2.3.4 Analysis of data with one observed fraction at each dose group 58
6.2.3.5 Extrapolation and ECx 58
6.2.3.6 Confidence intervals 58
6.2.4 Assumptions 59
6.2.4.1 General 59
6.2.4.2 Statistical assumptions 59
6.2.4.3 Evaluation of assumptions 59
6.2.4.3.1 Evaluation of basic assumptions 59
6.2.4.3.2 Evaluation of the additional assumption 59
6.2.4.4 Consequences of violating the assumptions 60
6.2.4.4.1 Consequences of violating basic assumptions 60
6.2.4.4.2 Consequences of violating the additional assumption 60
6.3 Dose-response modelling of continuous data (for a single exposure duration) 60
6.3.1 Purpose 60
6.3.2 Terms and notation 60
6.3.3 Choice of model 61
6.3.3.1 First distinctions 61
6.3.3.2 Linear models 62
6.3.3.3 Threshold models 62
6.3.3.4 Additive versus multiplicative models 63
6.3.3.5 Models based on "quantal" models 63
6.3.3.6 Nested non-linear models 64
6.3.3.7 Hill model 67
6.3.3.8 Non-monotone models 67
6.3.4 Model fitting and estimation of parameters 68
6.3.4.1 Software and assumptions 68
6.3.4.2 Response in controls 68
6.3.4.3 Fitting the model assuming normal variation 68
6.3.4.4 Fitting the model assuming normal variation after log-transformation 68
6.3.4.5 Fitting the model assuming normal variation after other transformations 69
6.3.4.6 No individual data available 69
6.3.4.7 Fitting the model using GLM 69
6.3.4.8 Covariates 70
6.3.4.9 Heterogeneity and weighted analysis 71
6.3.4.10 Confidence intervals 73
6.3.4.11 Extrapolation 73
6.3.4.12 Analysis of data with replicated dose group 73
6.3.5 Assumptions 74
6.3.5.1 General 74
6.3.5.2 Statistical assumptions 74
6.3.5.3 Additional assumption 74
6.3.6 Evaluation of assumptions 75
6.3.7 Consequences of violating the assumptions 75
6.3.7.1 Basic assumptions 75
6.3.7.2 Additional assumption 76
6.4 To accept or not accept the fitted model? 77
6.4.1 Can the fitted model be accepted and used for its intended purpose? 77
6.4.2 Is the model in agreement with the data? 77
6.4.3 Do the data provide sufficient information for fixing the model? 77
6.5 Design issues 81
6.5.1 General 81

6.5.2 Location of dose groups 81
6.5.3 Number of replicates 81
6.5.4 Balanced versus unbalanced designs 82
6.6 Exposure duration and time 82
6.6.1 General 82
6.6.2 Quantal data 82
6.6.3 Continuous data 83
6.6.3.1 General 83
6.6.3.2 Independent observations in time 83
6.6.3.3 Dependent observations in time 85
6.7 Search algorithms and non-linear regression 85
6.8 Reporting statistics 86
6.8.1 Quantal data 86
6.8.2 Continuous data 87

7 Biology-based methods 87
7.1 Introduction 87
7.1.1 Effects as functions of concentration and exposure time 87
7.1.2 Parameter estimation 89
7.1.3 Outlook 89
7.2 Modules of effect-models 90
7.2.1 General 90
7.2.2 Toxico-kinetic model 91
7.2.3 Physiological targets of toxicants 91
7.2.4 Change in target parameter 92
7.2.5 Change in endpoint 93
7.3 Survival 93
7.3.1 Relationship between hazard rate and survival probability 93
7.3.2 Assumptions of survival probability at any concentration of test compound 94
7.3.3 Summary 94
7.4 Body growth 97
7.4.1 Routes for affecting body growth 97
7.4.2 Assumptions 97
7.4.3 Von Bertalanffy growth curve 98
7.5 Reproduction 99
7.5.1 Routes that affect reproduction 99
7.5.2 Assumptions 100
7.5.3 Implication 100
7.6 Population growth 101
7.6.1 General 101
7.6.2 Assumptions 101
7.7 Parameters of effect models 103
7.7.1 General 103
7.7.2 Effect parameters 103
7.7.2.1 Toxicity and dynamic parameters 103
7.7.2.2 Killing rate, bk 104
7.7.3 Discussion 105
7.7.4 Eco-physiological parameters 107
7.8 Recommendations 109
7.8.1 Goodness of fit 109
7.8.2 Choice of modes of action 110
7.8.3 Experimental design 110
7.8.4 Building a database for raw data 110
7.9 Software support 110
7.9.1 General 110
7.9.2 DEBtox 111
7.9.3 DEBtool 111

8 List of existing guidelines with references to the subclauses of this Technical Specification 112

Annex A (informative) Analysis of an "acute immobilization of Daphnia magna" data set (OECD GL 202 -- ISO 6341) using the three presented approaches 115

A.1 Data set (see Table A.1) 115
A.2 Examples of data analysis using hypothesis testing (NOEC determination) 115
A.3 Example of data analysis by dose-response modelling 120
A.4 Example of data analysis using DEBtox (biological methods) 125

Annex B (informative) Analysis of an "algae growth inhibition" data set using the three presented approaches 127

B.1 General 127
B.2 Examples of data analysis using hypothesis testing (NOEC determination) 128
B.3 Example of data analysis by dose-response modelling 135
B.4 Examples of data analysis using DEBtox (biological methods) 139

Annex C (informative) Analysis of an "Daphnia magna reproduction" data set (OECD GL 211 - ISO 10706) using the three presented approaches 142

C.1 Examples of data analysis using hypothesis testing (NOEC determination) 143
C.2 Example of data analysis by dose-response modelling 148
C.3 Examples of data analysis using DEBtox (biological methods) 155

Annex D (informative) Analysis of a "fish growth" data set (OECD GL 204/215 - ISO 10229) using the three presented approaches 160

D.1 Data set 160
D.2 Examples of data analysis using hypothesis testing (NOEC determination) 162
D.3 Example of data analysis by dose-response modelling 172
D.4 Examples of data analysis using DEBtox (biological methods) 177

Annex E (informative) Description and power of selected tests and methods 180

E.1 Description of selected methods for use with quantal data 180
E.2 Power of the Cochran-Armitage test 189
E.3 Description of selected tests for use with continuous data 198
E.4 Power of step-down Jonckheere-Terpstra test 218

Annex F (informative) Annex to Clause 7 "Biology-based methods" 231

F.1 General 231
F.2 Effects on survival 231

Bibliography 237

Figure 1 -- Conceptual illustration of accuracy and precision 2

Figure 2 -- Illustration of a concentration-response relationship and of the estimates of the ECx and NOEC/LOEC 5

Figure 3 -- Analysis of quantal data: Methods for determining the NOEC 23

Figure 4 -- Analysis of continuous data: Methods for determining the NOEC 24

Figure 5 -- Analysis of continuous data: Methods for determining the NOEC (continued) 24

Figure 6 -- Flow-chart for dose-response modelling 50

Figure 7 -- Probit model fitted to observed mortality frequencies (triangles) as a function of log-dose 52

Figure 8 -- Logit model fitted to mortality dose-response data (triangles) 53

Figure 9 -- Weibull model fitted to mortality dose-response data (triangles) 54

Figure 10 -- Logit model fitted to mortality dose-response data (triangles), with background mortality 57

Figure 11 -- Two members from a nested family of models fitted to the same data set 66

Figure 12 -- Cholinesterase inhibition as a function of dose at three exposure durations 71

Figure 13 -- Relative liver masses against dose, plotted on log-scale 72

Figure 14 -- Dose-response model fitted to the data of Figure 13, showing that the heterogeneous variance was caused by males (triangles) and females (circles) responding differently to the chemical 73

Figure 15 -- Model fitted to dose-response data with and without an outlier in the top dose 76

Figure 16 -- Two different models (both with four parameters) fitted to the same data set resulting in similar dose-response relationships 79

Figure 17 -- Two data sets illustrating that passing a goodness-of-fit test is not sufficient for accepting the model 80

Figure 18 -- Observed biomasses (marks) as a function of time, for nine different concentrations of Atrazine 84

Figure 19 -- Growth rates as derived from biomasses observed in time (see Figure 18) at nine different concentrations (including zero), with the Hill model fitted to them 84

Figure 20 -- Estimated growth rates as a function of (log-)concentration Atrazine 85

Figure 21 -- Fluxes of material and energy through an animal, as specified in the DEB model 92

Figure 22 -- Time and concentration profiles of the hazard model, together with the data of Figure 27 95

Figure 23 -- Time and concentration profiles for effects on growth of Pimephalus promelas via an increase of specific maintenance costs by sodium pentachlorophenate (data by Ria Hooftman, TNO-Delft) 98

Figure 24 -- Time and concentration profiles for effects on growth of Lumbricus rubellus via a decrease of assimilation by copper chloride (data from Klok and de Roos 1996) 99

Figure 25 -- Effects of cadmium on the reproduction of Daphnia magna through an increase of the costs per offspring -- Data from the OECD ring-test 101

Figure 26 -- Example of application of the DEBtox method 102

Figure 27 -- The effect of a mixture of C,N,S-compounds on the growth of Skeletonema costatum via an increase of the costs for growth -- Data from the OECD ring test 103

Figure 28 -- A typical table of data that serves as input for the survival model, as can be used in the software package DEBtox (Kooijman and Bedaux 1996) 108

Figure 29 -- This profile likelihood function of the NEC (right panel) for the data in Figure 28 results from the software package DEBtox (Kooijman and Bedaux 1996) 108

Figure A.1 -- Probit model fitted to mortality response at day 2 -- CED = EC10 122

Figure A.2 -- The Weibull (left panel) and the two-stage LMS model fitted to the mortality data at day 2 123

Figure A.3 -- Probit model fitted to mortality data on day 1 (left panel), and fitted to both day 1 and day 2 simultaneously 124

Figure A.4 -- DEBtox example: Parameters and asymptotic standard deviations (ASD) 125

Figure A.5 -- Graphical test of model predictions against data 125

Figure B.1 -- Exponential growth model fitted to biomass, assuming a constant initial biomass (a), and growth rate (b) dependent on concentration (0, 0,01, 0,02, 0,03, 0,06, 0,1, 0,2, 0,3, 0,6 mg/l) 136

Figure B.2 -- Exponential growth model fitted to biomass, assuming a constant initial biomass (a), and growth rate (b) dependent on each individual flask (six for the control group, and 2 for each nonzero concentration) 136

Figure B.3 -- Estimated growth rates (from individual flasks, see Figure B.2) as a function of the concentration of Atrazine 137

Figure B.4 -- Regression residuals from analysis on log-scale (upper panels) and from analysis without transformation (lower panels) 138

Figure B.5 -- Growth rate plots 139

Figure B.6 -- DEBtox example: Data for effects of Atrazine in micrograms per litre on the growth of Selenastrum capricornutum in cells per millilitre 139

Figure B.7 -- DEBtox example: Parameter estimates and asymptotic standard deviations (ASD) 139

Figure B.8 -- DEBtox example: Time profile (Population growth, growth model, CAS 1912-24-9) 140

Figure B.9 -- DEBtox example: Concentration profile (Population growth, growth model, CAS 1912-24-9) . 140 Figure B.10 -- DEBtox example: Profile likelihood for NEC estimate (Population growth, growth model, CAS 1912-24-9) 141

Figure C.1 -- Number of life young as a function of concentration (on log-scale to improve visibility), counted over the first two weeks (triangles) and over the third week (circles) .149

Figure C.2 -- Total live young (TLY) in third week plotted against TLY in first two weeks, showing the correlations between these counts 149

Figure C.3 -- Exponential model fitted to the number of life young counted over week three 151

Figure C.4 -- Plots of regression residuals 152

Figure C.5 -- Means of number of young, plotted cumulatively against time 153

Figure C.6 -- Estimated ET50 values from Figure C.5, plotted against the concentration, with a fitted dose-response model 154

Figure C.7 -- Number of young, plotted cumulatively against time 154

Figure C.8 -- ET50s estimated per replicate (see Figure C.7) as a function of the concentration with a fitted dose-response model 155

Figure C.9 -- ET50s estimated per replicate (see Figure C.7) as a function of the concentration, with two outliers removed 155

Figure C.10 -- DEBtox example: Data for the cumulative number of offspring per female as affected by an unknown compound 156

Figure C.11 -- DEBtox example: Parameter estimates and asymptotic standard deviations (ASD) ...156

Figure C.12 -- DEBtox example: Time profile (reproduction, maintenance model, ISO repro set) 157

Figure C.13 -- DEBtox example: Concentration profile (reproduction, maintenance model, ISO repro set) . 157 Figure C.14 -- DEBtox example: Profile likelihood for NEC estimate (reproduction, maintenance model, ISO repro set) 158

Figure C.15 -- DEBtox example: Body length at 21 days 158

Figure D.1 -- Exponential model, y = a exp(bx), fitted to masses at 28 days 174

Figure D.2 -- Plots of regression residuals for the analysis of Figure D.1 175

Figure D.3 -- Plots of regression residuals 175

Figure D.4 -- Dose-response analysis of the fish masses, but without log-transformation 176

Figure D.5 -- Exponential model fitted to the body lengths 177

Figure D.6 -- DEBtox example: Parameter estimates and asymptotic standard deviations (ASD) 177

Figure D.7 -- DEBtox example: Concentration profile 178

Figure D.8 -- DEBtox example: Profile likelihood for NEC estimate 179

Figure E.1 -- Cochran-Armitage test: Plot showing that 5 subjects per concentration would give very low power 190

Figure E.2 -- Cochran-Armitage test: Design with 20 subjects per concentration 191

Figure E.3 -- Cochran-Armitage power versus maximum rate change: Power at Dose 5 in 5-dose study, with trend shape linear, lag = 0, sample size = 20 192

Figure E.4 -- Cochran-Armitage power versus maximum rate change: Power at Dose 4 in 5-dose study, with trend shape linear, lag = 0, sample size = 20 192

Figure E.5 -- Cochran-Armitage power versus maximum rate change: Power at Dose 3 in 5-dose study, with trend shape linear, lag = 0, sample size = 20 193

Figure E.6 -- Cochran-Armitage power versus maximum rate change: Power at Dose 5 in 5-dose study with trend shape linear, lag = 0, sample size = 40 193

Figure E.7 -- Cochran-Armitage power versus maximum rate change: Power at Dose 4 in 5-dose study with trend shape linear, lag = 0, sample size = 40 194

Figure E.8 -- Cochran-Armitage power versus maximum rate change: Power at Dose 3 in 5-dose study with trend shape linear, lag = 0, sample size = 40 194

Figure E.9 -- Cochran-Armitage power versus maximum rate change: Power at Dose 5 in 5-dose study with trend shape linear, lag = 0, sample size = 60 195

Figure E.10 -- Cochran-Armitage power versus maximum rate change: Power at Dose 4 in 5-dose study with trend shape linear, lag = 0, sample size = 60 195

Figure E.11 -- Cochran-Armitage power versus maximum rate change: Power at Dose 3 in 5-dose study with trend shape linear, lag = 0, sample size = 60 196

Figure E.12 -- Cochran-Armitage power versus maximum rate change: Power at Dose 5 in 5-dose study with trend shape linear, lag = 0, sample size = 80 196

Figure E.13 -- Cochran-Armitage power versus maximum rate change: Power at Dose 4 in 5-dose study with trend shape linear, lag = 0, sample size = 80 197

Figure E.14 -- Cochran-Armitage power versus maximum rate change: Power at Dose 3 in 5-dose study with trend shape linear, lag = 0, sample size = 80 197

Figure E.15 -- Power of step-down Jonckheere test: 6 doses at Step 1, N = 10 219

Figure E.16 -- Power of step-down Jonckheere test: 6 doses at Step 2 versus 5 doses at Step 1, N = 10 220

Figure E.17 -- Power of step-down Jonckheere test: 6 doses at Step 3 versus 5 doses at Step 2 versus 4 doses at Step 1, N = 10 220

Figure E.18 -- Power of step-down Jonckheere test: 6 doses at Step 2 versus 5 doses at Step 1, N = 5 221

Figure E.19 -- Power of step-down Jonckheere test: 6 doses at Step 1, N = 5 221

Figure E.20 -- Power of step-down Jonckheere test: (Doses,Step) = (6,4) (5,3)(4,2) (3,1), N = 5 222

Figure E.21 -- Power of step-down Jonckheere test: 6 doses at Step 3 versus 5 doses at Step 2 versus 4 doses at Step 1, N = 5 222

Figure E.22 -- Power of step-down Jonckheere test: 6 doses at Step 1, N = 10 223

Figure E.23 -- Power of step-down Jonckheere test: 6 doses at Step 2 versus 5 doses at Step 1, N =10 223

Figure E.24 -- Power of step-down Jonckheere test: 6 doses at Step 3 versus 5 doses at Step 2 versus 4 doses at Step 1, N = 10 224

Figure E.25 -- Power of step-down Jonckheere test: (Doses, Step) = (6,4) (5,3) (4,2) (3,1), N = 10 224

Figure E.26 -- Power of step-down Jonckheere test: 6 doses at Step 1, N = 20 225

Figure E.27 -- Power of step-down Jonckheere test: 6 doses at Step 2 versus 5 doses at Step 1, N = 20 225

Figure E.28 -- Power of step-down Jonckheere test: 6 doses at Step 3 versus 5 doses at Step 2 versus 4 doses at Step 1, N = 20 226

Figure E.29 -- Power of step-down Jonckheere test: (Doses, Step) = (6,4) (5,3) (4,2) (3,1), N = 20 226

Figure E.30 -- Power of step-down Jonckheere test: 6 doses at Step 1, N = 40 227

Figure E.31 -- Power of step-down Jonckheere test: 6 doses at Step 2 versus 5 doses at Step 1, N = 40 227

Figure E.32 -- Power of step-down Jonckheere test: 6 doses at Step 3 versus 5 doses at Step 2 versus 4 doses at Step 1, N = 40 228

Figure E.33 -- Power of step-down Jonckheere test: (Doses, Step) = (6,4) (5,3) (4,2) (3,1), N = 40 228

Figure E.34 -- Power of step-down Jonckheere test: 6 doses at Step 1, N = 80 229

Figure E.35 -- Power of step-down Jonckheere test: 6 doses at Step 2 versus 5 doses at Step, N = 80 229

Figure E.36 -- Power of step-down Jonckheere test: 6 doses at Step 3 versus 5 doses at Step 2 versus 4 doses at Step 1, N = 80 230

Figure E.37 -- Power of step-down Jonckheere test: (Doses, Step) = (6,4) (5,3) (4,2) (3,1), N = 80 230