EBCTCG seventh cycle variables and data format

Either using the codes we suggest below or using your own codes, please extract from your dataset the variables that correspond most closely to the items listed below and send them to us. Please provide one record for each person ever randomised (including any person who was randomised and then was later categorised as ineligible, withdrawn, unevaluable, lost or "protocol deviant"—but, please tell us in question 10 which patients your group's preferred analyses would exclude, and why).

For trials where a dataset has previously been sent to the EBCTCG it is probably easiest and most reliable to update by re-sending all variables. If, however, this would cause difficulties then you can send only the additional variables; let us know if you want a file of the data you previously sent and we will provide it.

If any variable is not available or not applicable, please omit it and send only the remaining variables. If you have any of the requested variables in your records in a form that would require substantial additional work to supply (e.g. computerisation, or manual coding), feel free to omit them for now, but in your cover document please tell us of their existence. Please send the following:

- Your data in a separate Excel spreadsheet for each separate trial, if possible.
- A cover document giving all your coding conventions (including your format for dates).
- Send data to: bc.overview@ndph.ox.ac.uk with your research group's name (and/or the EBCTCG number for your research group) and your group's name for the trial in the subject line.

If you have any questions about this data request, please contact the EBCTCG secretariat on bc.overview@ndph.ox.ac.uk (Telephone: +44-1865-743852). All data supplied to the secretariat will be held securely and treated confidentially, in accordance with the EBCTCG Data Policy (available at: https://www.ctsu.ox.ac.uk/research/ebctcg).

For additional notes on supplying data to the overview, please see the EBCTCG Collaborators' Space web pages, or email the EBCTCG secretariat

CORE VARIABLES—BASELINE (Q1–29)

A) Randomisation and patient characteristics (Q1–11)

1. **Your patient identifier** *(preferably specifying uniquely which trial as well as which patient)*
2. **Date of randomisation** *(specify your format for dates [in your covering document])*
3. **Allocated treatment** *(specify your codes)*
4. **Age** at randomisation *(years)* **NB Here & everywhere else, leaving an item Blank means Not Known**
5. **Height** at randomisation *(m)*
6. **Weight** at randomisation *(kg)*
7. **Ethnicity/Race** *(if available: please use your own group's preferred method of coding)*
8. **Country of origin of patient, or geographical region** *(if available: please use your own group's preferred method of coding)*
9. **Menopausal status** at randomisation *(1=pre-, 2=peri-, 3=postmenopausal with intact ovaries & uterus, 4=ovarian ablation, 5=hysterectomy, 6=both [ie, 4 and 5]), 7=artificial, 8=male patient)*
10. **Did chemotherapy cause apparently permanent cessation of menses?** *(1=no/not applicable, 2=yes)*
11. **Would your group's preferred analyses exclude this patient?** NB A few trial patients may be randomised in error, otherwise ineligible, lost with no follow-up, unevaluable or withdraw consent. *(1=no known reason for exclusion, 2=yes [specify main reason(s) for preferring exclusion, if known])*

B) Surgical details (Q12-13; or, define and use your own codes)

12. **Breast surgery** *(1=none, 2=only lumpectomy or wide local excision, 3=quadrantectomy/sector resection, 4=partial mastectomy, 5=simple/total mastectomy, 6=radical mastectomy, 7=modified radical mastectomy)*
13. **Axillary surgery** *(1=none, 2=sentinel node biopsy only, 3=axillary sampling, 4=surgical clearance of less than levels I & II, 5=full clearance of axillary levels I & II, 6=clearance of more than levels I & II, 7=axillary clearance, but levels cleared unspecified)*

C) Pathological nodal status (Q14–15; or, use your own codes [eg, TNM])

Note: In patients receiving neo-adjuvant treatment (or in trials where neo-adjuvant or axillary treatment differs between groups) give nodal status prior to neo-adjuvant (or axillary) therapy in section K

14. **Sentinel node biopsy** *(1=not done; 2=done and negative for cancer; 3=no greater involvement than isolated tumour cells [≤0.2 mm and/or < 200 cells]; 4=no greater involvement than micrometastases [>0.2mm or >200cells but ≤2mm]; 5=macroscopic nodal deposits [>2 mm]; 6=positive but size unknown)*
15. **Axillary status** *(specify codes, or: 1=pN- histologically; 2=N- other/unknown method; 3=1-3 positive nodes ; 4=4-9 [or 4+] positive; 5=10+ positive; 6=N+ histologically, unknown number; 7=N+ other/unknown method)*

D) Tumour characteristics (Q16–20; or, use your own codes [eg, TNM])

16. **Method first detected** *(1=mammographic screening, 2=incidental, 3=symptomatic, 4=other)*
17. **Laterality** *(1=left, 2=right, 3=bilateral)*
18. **Pathological grade prior to any neo-adjuvant therapy** *(1=well differentiated, 2=moderately,3=poorly)*
19. **Histological type** (if not locally determined please state)*(1=invasive, not otherwise specified, 2=ductal, 3=lobular, 4=other invasive, 5=mixed, 6=carcinoma in situ (CIS) only)*
20. **Tumour diameter:** largest diameter of excised primary *(mm)*

E) Receptor status (Q21–29; or, use your own codes)

Note: In trials with some neo-adjuvant treatment give receptor status prior to any neo-adjuvant therapy

21. **Summary of Estrogen Receptor (ER) status of primary tumour** *(1=ER-poor, 2=ER+, 3=ER++ [define in cover document, unless ER-poor is <10 fMol/mg and ER++ is ER definitely ≥100 fMol/mg])*
22. **Quantitative ER measurement** *(measured in central/reference lab if possible, otherwise best available)*
23. **Units for ER** *(1=fMol/mg, 2=% +ve by IHC, 3=Allred score [category score], 4=H-score, 9=other [specify]*)
24. **Summary of Progesterone Receptor (PR) status of primary tumour** *(1=PR-poor, 2=PR+, 3=PR++ [define in cover document, unless PR-poor is <10 fMol/mg and PR++ is PR definitely ≥100 fMol/mg])*
25. **Quantitative PR measurement** *(done in central/reference lab if possible, otherwise best available)*
26. **Units for PR** *(coded as Q21)*
27. **Summary of HER2 status of primary** *(1=negative/normal, 2=positive/over-expressing)*
28. **Quantitative HER2 measurement** *(done in central/reference lab if possible, otherwise best available)*
29. **Units for HER2** *(1=IHC [% staining], 2=IHC score [0, 1+, 2+, 3+], 3=FISH [# copies], 4=FISH [HER2:CEP17 ratio], 5=CISH [# copies], 6=CISH [HER2:CEP17], 9=other [please specify])*

CORE VARIABLES—FOLLOW-UP (Q30-45)

F) Non-compliance before any recurrence (Q30-31; or, use your own codes)

30. **Any substantial deviation from trial treatment allocation (before any breast cancer recurrence)?** *(1=no, 2=never started, 3=discontinued, 4=switched to opposite trial group, 5=other [specify])*

31. **Date of first such deviation from allocated treatment** *(ignore deviations after recurrence)*

G) Cancer recurrence and second cancers (Q32–42; or, use your own codes)

32. **Any recurrence of invasive breast cancer (ie, locoregional, contralateral or distant)?** NB Includes any occurrence of new ipsilateral or contralateral breast cancer *(1=no, 2=yes)*

33. **<u>If no:</u> Date patient last known to be free of such recurrence; <u>If yes:</u> Date of first such recurrence**

34. **Site of first distant recurrence (ie, possibly distant; not just locoregional/contralateral)** *(1=no distant recurrence, 2=recurrence-unknown if distant, 3=distant recurrence-unknown site(s), 4=only in distant soft tissue, 5=only in distant nodes, 6=only in bone, 7=only visceral, 8=only in CNS, 9=multiple sites including bone but not CNS/brain, 10=multiple sites not including bone or CNS/brain, 11=multiple sites including CNS/brain but not bone, 12=multiple sites including CNS/brain and bone)*

35. **Date of first distant recurrence** NB Locoregional recurrence can precede first distant recurrence

36. **Site of first locoregional recurrence** *(1=no locoregional recurrence recorded, 2=multiple or unspecified locoregional sites 3=only in breast [new or recurrent invasive cancer] or chest wall, 4=only in axilla, 5=only in other locoregional nodes [eg, infraclavicular fossa], 6= multiple locoregional sites, 7=only in internal mammary nodes, 8=only in supraclavicular nodes, 9=tumour bed, 10=breast but known not tumour bed, 11=CIS only (if index CIS or if subsequent invasive cancer not collected))*

37. **Date of first locoregional recurrence**

38. **Contralateral breast cancer?** *(1=no, 2=yes: new invasive cancer thought to have arisen during follow-up in the contralateral breast, 3=CIS only (if index CIS))*

39. **Date of first contralateral breast cancer**
NB If patient had more than one second malignancy during follow-up, <u>repeat</u> variables 39-41 for each.

40. **Site of any second malignancy *[except breast cancer (including contralateral)]* during follow-up** *(Describe ALL sites. Use and specify your own codes; if you use ICD codes specify revision, eg ICD-9 or ICD-10)*

41. **Date of this second malignancy**

42. **MIGHT this have been a breast cancer metastasis?** *(1=no, 2=possibly/not yet certain [eg, possible lung, liver, bone or brain metastasis: please do not report definite breast metastases as second cancers])*

H) Survival (Q43−45)

43. **Is patient known to have died?** *(1=no, 2=yes)*

44. **<u>If NO:</u> Date patient last known to be alive; <u>If yes:</u> Date of death**

45. **<u>If YES:</u> Cause of death** *(use and specify your own codes; if you use ICD codes specify which version, eg ICD-9 or ICD-10)*

ADDITIONAL VARIABLES (Q46–61)

I) Additional tumour marker data (Q46–54; or, use your own codes)

Note: If tests of gene expression or special tests of IHC quantitation were done on the excised primary then please send a separate file in your own format with the fully detailed set of results on each individual.

46. **Summary of gene-expression status of primary tumour** *(1=low risk, 2=intermediate risk, 3=high risk): NB Please also provide the fully detailed gene expression results for each patient as a separate dataset.*
47. **Quantitative gene-expression prognostic score** *(best available single numerical measure)*
48. **Prognostic score used to quantify gene expression profile** *(use own code, or: 1=OncotypeDx prognostic score, 2=Mammaprint prognostic score, 3=EndoPredict, 4=Prosigna,9=other [please specify])*

49. **Summary of Topo-isomerase II alpha (TOPO2A) status of primary tumour** *(1= normal [ie, no gene over-expression or deletion], 2=positive/over-expressing, 3=deleted)*
50. **Quantitative TOPO2A measurement** *(done in central/reference laboratory if possible)*
51. **Units for TOPO2A** *(1=IHC [% staining], 2=IHC score [0, 1+, 2+, 3+], 3=FISH [number of copies], 4=FISH [TOPO:CEP17 ratio], 5=CISH [# copies], 6=CISH [TOPO:CEP17], 9=other [please specify])*

52. **Summary of Proliferation Index of primary tumour** *(1=low, 2=intermediate, 3=high)*
53. **Quantitative Proliferation Measure** *(best available numerical measure, in central/ ref lab if possible)*
54. **Factor measured for Proliferation Index** *(1=S-phase fraction [%], 2=thymidine labelling index [%], 3=Ki-67 by IHC [% staining], 9=other [please specify]*)

J) Non-fatal adverse events (Q55–56; omit if not sought)

Note: Some treatments may cause or prevent bone fractures, cardiovascular events, lymphoedema, or lung fibrosis. Please describe all such events (eg, hip fracture, spinal fracture, myocardial infarction, stroke, pulmonary embolus, episode of cardiac failure) if, but only if, such events were sought and recorded systematically for all arms of the trial.

If more than one such event was recorded, repeat variables 54-55 for each.
55. **Nature and severity of event** *(use your own codes; if you use ICD codes, specify which version, eg ICD-9 or ICD-10, and if you use CTC Adverse Event codes, please specify version number, eg CTCAE-3 or CTCAE-4)*
56. **Date of event**

K) Trials with some neo-adjuvant systemic therapy or where axillary treatment differs between groups (Q57–61; or, use own codes)

57. **Apparent axillary nodal status (clinical, radiological or other) before neo-adjuvant** *(1=N-, 2=N+)*
58. **Apparent tumour diameter (clinical or radiological) before neo-adjuvant***:* largest diameter *(mm)*
59. **Operability before any neo-adjuvant therapy** *(define your own codes, or: 1=Breast-conserving surgery feasible, 2=Mastectomy but not BCS feasible, 3=inoperable, 4=uncertain operability)*
60. **Breast tumour response after completion of neo-adjuvant** *(define your own codes, or: 1=clinically complete response [cCR] & negative pathology (for invasive disease and DCIS), 2=cCR with DCIS, 3=cCR with invasive cancer remaining pathologically, 4=cCR with no pathological information, 5=partial response, 6=stable disease, 7=progression [define 5–7])*
61. **Axillary response after neo-adjuvant** *(coded as Q59)*

ADDITIONAL VARIABLES FOR SPECIFIC META-ANALYSES (Q62–93)

For some meta-analyses we may need additional information on surgery, radiotherapy, adjuvant treatments received, or additional tumour markers but this section does not need to be completed for most meta-analyses so omit unless specifically requested.

L) Trials of extended endocrine therapy (Q62–65)

Note: This information is needed only for trials of longer versus shorter endocrine therapy

62. **Endocrine therapy given prior to randomisation** *(1=Tamoxifen, 2=Aromatase Inhibitor, 3=Tamoxifen then Aromatase Inhibitor, 4=Aromatase Inhibitor then Tamoxifen, 5=Other (specify))*
63. **Date initial endocrine therapy started** *(Approximate date, or date of surgery if unknown)*
64. **Date of first switch from tamoxifen to AI, or vice versa** *(if applicable)*
65. **Date initial endocrine therapy completed** *(Approximate date)*

M) Details of additional treatments for trials of local, biological, or endocrine therapy (Q66–70)

66. **Neoadjuvant chemotherapy received** *(1=no, 2=non-anthracycline, non-taxane, 3=anthracycline, non-taxane, 4=taxane+anthracycline, 5=other taxane-containing, 9=Yes – type unknown)*
67. **Adjuvant chemotherapy received** *(as for Q66)*
68. **Endocrine therapy received** *(1=no, 2=tamoxifen, 3=aromatase inhibitor, 4=sequential tamoxifen and aromatase inhibitor, 5=ovarian ablation/suppression alone, 6=ovarian ablation+tamoxifen, 7=ovarian ablation + aromatase inhibitor, 9=Yes – type unknown))*
69. **HER2 directed therapy received** *(1=no, 2=Yes)*
70. **Radiotherapy received (irrespective of site)** *(1=No, 2=Yes)*

N) Additional therapy details for trials of local therapy (Q71–81)

71. **Date of first breast surgery** *(dd/mm/yyyy)*
72. **Site of tumour in breast quadrant** *(1=lateral, 2=medial, 3=central, 4=medial or central, 5=not specified)*
73. **Lymphovascular invasion** *(1=no, 2=yes)*
74. **Date of first axillary surgery** *(dd/mm/yyyy)*
75. **Total number of sentinel lymph nodes excised and examined pathologically** *(-1=none, 1=one, 2=two, etc)*
76. **Number of sentinel lymph nodes excised and were:**
    a. **isolated tumour cells [≤0.2 mm and or ≤200 cells]** *(-1=none, 1=one, 2=two, etc)*
    b. **micrometastasis (>0.2mm or >200cells but ≤2mm)** *(-1=none, 1=one, 2=two, etc)*
    c. **macroscopic nodal deposits [>2 mm]** *(-1=none, 1=one, 2=two, etc)*
    d. **positive (unknown whether ITC, micro or macrometastasis)** *(-1=none, 1=one, 2=two, etc)*
    e. **negative** *(-1=none, 1=one, 2=two, etc)*

77. **Radiotherapy to whole breast/chest wall** *(1=No, 2=Yes)*
78. **Radiotherapy to partial breast** *(1=No, 2=Yes)*
79. **Radiotherapy to supraclavicular fossa** *(1=No, 2=Yes)*
80. **Radiotherapy to axilla** *(1=No, 2=Yes)*
81. **Radiotherapy to internal mammary chain** *(1=No, 2=Yes)*

O) Additional details for trials of DCIS (Q82–86)

82. **Closest relevant excision margin** *(please give in mm where available; 99=no excision performed)*
83. **Diagnosis at entry** *(1=DCIS only, 2=DCIS+LCIS, 3 = DCIS (+/- LCIS) with microinvasion, 4=unknown)*
84. **Comedo** *(1=Present, marked or severe; 2=Present, moderate; 3=Present, slight; 4=Present, NOS; 5=Absent; 6=Unknown)*
85. **Architecture** *(1=Cribriform; 2=Micropapillary; 3=Papillary; 4=Solid; 5=Other (please specify), 6=Unknown)*
86. **Focality** *(1=Unifocal;2=Multifocal/multicentric;3=Unknown)*

P) Additional questions for trials of duration of biological therapy *(Q87–89)*

Note: This information is needed only for trials of longer versus shorter biological therapy (e.g. trastuzumab duration)

87. Date initial biological therapy started *(Approximate date, or date of surgery if unknown)*
88. Date initial biological therapy completed *(Approximate date)*
89. Biological therapy given prior to randomisation *(1=None, 2=Yes)*

Q) Further tumour markers *(Q90–93)*

90. Summary of Tumour Infiltrating Lymphocytes (TILS) status *(1=low, 2=intermediate, 3=high)*
91. Quantitative TILS Measure *(% staining positive, in central/ ref lab if possible)*
92. Central histological type *(coded as for Q18)*
93. E-Cadherin data *(1=negative, 2=positive,3=intermediate/unclear)*