EBCTCG seventh cycle, 2016-20 variables and data format for trials of systemic therapy

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for each woman ever randomised (including any woman who was randomised and then was later categorised as ineligible, withdrawn, unevaluable, lost or "protocol deviant" – but, please tell us in question 9 which patients your group's preferred analyses would exclude, and why).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), please feel free to omit them for now, but in your cover document please tell us of their existence.

- Please send your data in a separate Excel spreadsheet for each separate trial, if possible.
- Please send a cover document giving all your coding conventions (including your format for dates).
- Please send your 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 (tel: +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) and the UK Medical Research Council data security policies.

CORE VARIABLES - BASELINE (Q1-27)

A) Randomisation and patient characteristics (Q1-9)

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. **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])*
8. **Did chemotherapy cause apparently permanent cessation of menses?** *(1=no/not applicable, 2=yes)*
9. **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 (Q10-11; or, define and use your own codes)

10. **Breast surgery** *(1=none, 2=only lumpectomy or wide local excision, 3=quadrantectomy or sector resection, 4=partial mastectomy, 5=simple or total mastectomy, 6=radical mastectomy)*
11. **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) Nodal status (Q12-13; or, use your own codes [eg, TNM])

12. **Sentinel node biopsy** *(1=not done, 2=done and negative for cancer, 3= isolated tumour cells [≤0.2 mm], 4=micrometastasis (0.2 to 2mm), 5=macroscopic nodal deposit [>2 mm], 6=positive, size unknown)*
13. **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 (Q14-18; or, use your own codes [eg, TNM])

14. **Method first detected** *(1=mammographic screening, 2=incidental, 3=symptomatic, 4=other)*
15. **Laterality** *(1=left, 2=right, 3=bilateral)*
16. **Pathological grade** *(1=well differentiated, 2=moderately,3=poorly)*
17. **Histological type** *(1=invasive, not otherwise specified, 2=ductal, 3=lobular, 4=other, 5=mixed)*
18. **Tumour diameter:** largest diameter of excised primary *(mm)*

E) Receptor status (Q19-27; or, use your own codes)

Note: In trials with some neo-adjuvant systemic therapy give receptor status at randomisation, i.e. prior to any neo-adjuvant therapy

19. **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 fm/mg and ER++ is ER definitely ≥100 fm/mg])*
20. **Quantitative ER measurement** *(measured in central/reference lab if possible, otherwise best available)*
21. **Units for ER** *(1=fm/mg, 2=% +ve by IHC, 3=Allred score [category score], 4=H-score, 9=other [specify])*

22. **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 fm/mg and PR++ is PR definitely ≥100 fm/mg])*
23. **Quantitative PR measurement** *(done in central/reference lab if possible, otherwise best available)*
24. **Units for PR** *(coded as Q21)*

25. **Summary of HER2 status of primary** *(1=negative/normal, 2=positive/over-expressing)*
26. **Quantitative HER2 measurement** *(done in central/reference lab if possible, otherwise best available)*
27. **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 (Q28-43)

F) Non-compliance before any recurrence (Q28-29; or, use your own codes)

28. 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])*

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

G) Cancer recurrence and second cancers (Q30-40; or, use your own codes)

30. 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)*

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

32. 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/multiple sites, 4=only in distant soft tissue, 5=only in distant nodes, 6=only in bone, 7=only visceral, 8=only in CNS)

33. Date of first distant recurrence NB Locoregional recurrence can precede first distant recurrence

34. 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, supraclavicular or internal mammary])*

35. Date of first locoregional recurrence

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

37. Date of first contralateral breast cancer

NB If patient had more than one second malignancy during follow-up, <u>repeat</u> variables 38-40 for each.

38. Site of any second malignancy *[except breast cancer]* 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)*

39. Date of this second malignancy

40. 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 (Q41-43)

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

42. <u>If no:</u> Date patient last known to be alive; <u>If yes:</u> Date of death

43. <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 (Q44-61)

I) Additional tumour marker data (Q44-52; 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.

44. **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.
45. **Quantitative gene-expression prognostic score** *(best available single numerical measure)*
46. **Prognostic score used to quantify gene expression profile** *(use own code, or: 1=OncotypeDx prognostic score, 2=Mammaprint prognostic score, 9=other [please specify])*

47. **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)*
48. **Quantitative TOPO2A measurement** *(done in central/reference laboratory if possible)*
49. **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])*

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

J) Bone fractures and cardiovascular events (Q53-54; omit if not sought)

Some trial treatments may cause or prevent bone fractures or cardiovascular events. 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 the trial.

If more than one such event was recorded, repeat variables 53-54 for each.

53. **Nature 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)*
54. **Date of event**

K) Trials with some neo-adjuvant systemic therapy (Q55-59; or, use own codes)

55. **Apparent axillary nodal status (clinical, radiological or other) before neo-adjuvant** *(1=N-, 2=N+)*
56. **Apparent tumour diameter before neo-adjuvant***:* largest diameter *(mm)*
57. **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)*
58. **Breast tumour response after neo-adjuvant** *(define your own codes, or: 1=clinically complete response [cCR] & negative pathology, 2=cCR with DCIS, 3=cCR, cancer remaining, 4=cCR with no pathological information, 5=partial response, 6=stable disease, 7=progression [define 5-7])*
59. **Axillary response after neo-adjuvant** *(coded as Q58)*

L) Trials of extended endocrine therapy (Q60-63)

60. **Prior endocrine therapy given** (1=Tamoxifen, 2=Aromatase Inhibitor, 3=Tamoxifen plus Aromatase Inhibitor)
61. **Date endocrine therapy started** (or date of surgery if unknown)
62. **Date initial endocrine therapy stopped** (ignoring any re-introduction after recurrence)
63. **Chemotherapy given** (1=No, 2=Yes, 3=Unknown)