**Supplementary Tables and Figures**

**Table S1.** Set-up of the search strategy used to identify articles for the systematic literature review

|  |
| --- |
| **Search aim:**Identify all publications reporting prognostic factors associated with subsequent ipsilateral invasive breast cancer after an initial diagnosis of DCIS. |
| **Mesh terms:**

|  |  |  |
| --- | --- | --- |
|  | **Question Concepts** | **Search Term(s) for PubMed** |
| **P** | BreastDCIS | \* Breast Neoplasms\* Carcinoma, Intraductal, Noninfiltrating\* Carcinoma in Situ |
| **I** | - |  |
| **C** | - |  |
| **O** | ProgressionRecurrence  | \* Disease Progression\* Neoplasm Recurrence, Local |
|  | Biomarkers, Prognostic/predictive markers, Predictors, candidate | \* Biomarkers, Tumor |

PICO method for PubMed search: P = Patients/population; I = Intervention (not appropriate); C = Comparison of intervention (not appropriate); O = Outcome we would like to measure |
| **Text words for “DCIS”:**

|  |  |
| --- | --- |
| “ductal carcinoma in situ”“DCIS”“intraductal carcinoma in situ”“pre-invasive ductal carcinoma\*” “preinvasive ductal carcinoma\*”“non-invasive ductal carcinoma\*”“noninvasive ductal carcinoma\*”“pre-invasive breast carcinoma\*”“preinvasive breast carcinoma\*”“non-invasive breast carcinoma\*”“noninvasive breast carcinoma\*”“pre-invasive breast tum\*”“preinvasive breast tum\*”“non-invasive breast tum\*”“noninvasive breast tum\*”“stage zero breast cancer\*” | “non-infiltrating intraductal carcinoma\*”“noninfiltrating intraductal carcinoma\*”“intra-ductal carcinoma\*”“intraductal carcinoma\*”“mammary intra-epithelial neoplasia\*”“mammary intraepithelial neoplasia\*”“ductal intra-epithelial neoplasia\*”“ductal neoplas\*”“non-invasive breast\*” “noninvasive breast\*”“pre-invasive breast tum\*”“preinvasive breast tum\*”“non-infiltrating carcinoma\*”“noninfiltrating carcinoma\*” |

 |

**Table S2.** Quality of prognosis studies (QUIPS) tool. Modified from Hayden *et al*. 2006.

|  |  |  |  |
| --- | --- | --- | --- |
| **Domains** | **Issues for consideration** | **Prompting questions** | **Ratings** |
| **Study participation** | 1. Population of interest
2. Method used to identify population
3. Recruitment period\*
4. Place of recruitment\*
5. Inclusion and exclusion criteria
6. Baseline characteristics
 | 1. Population described and selected without bias?
2. Is the source of medical records described?
3. Period of recruitment?\*
4. Place of recruitment?\*
5. Inclusion and exclusion (mastectomy, adjacent invasive disease, prior cancer diagnosis) criteria adequately and fully described?
6. Information given on age, grade (gr 1: 20-25%. gr 2:50%. gr 3: 25-30%; Additional: ER+ 80%, HER2+ 30%), clinical presentation, treatment, lesion size, margin status?
 | **High bias**: The relationship between the prognostic factor (PF) and outcome is very likely to be different for participants and eligible nonparticipants**Moderate bias**: The relationship between the PF and outcome may be different for participants and eligible nonparticipants**Low bias**: The relationship between the PF and outcome is unlikely to be different for participants and eligible nonparticipants |
| **Study attrition** | 1. Proportion of baseline samples available for analysis
2. Attempts to collect information on participants who dropped out
3. Reason and potential impact of loss to follow-up
4. Outcome and PF information on those who dropped out
 | 1. Adequate proportion of patients included in analysis?
2. Information collected on drop-outs?
3. Reason and impact of drop-out given?
4. Are the missing data at random or not at random? Important difference?
 | **High bias**: The relationship between the PF and outcome is very likely to be different for dropouts and non-dropouts**Moderate bias**: The relationship between the PF and outcome may be different for drop-outs and non-drop-outs**Low bias**: The relationship between the PF and outcome is unlikely to be different for dropouts and non-dropouts |

|  |  |  |  |
| --- | --- | --- | --- |
| **Domains** | **Issues for consideration** | **Prompting questions** | **Ratings** |
| **Prognostic factor measurement** | 1. Definition of the PF
2. Valid and reliable measurement of PF
3. Method and setting of PF
4. Proportion of data PF analysis
5. Method used for missing data
 | 1. Clear definition provided?
2. Variable described in both patients that did and did not experience an event of invasive breast cancer?
3. Method valid and reliable (blinded by outcome)? The same for all study participants? Continuous variables reported or appropriate cut-points used?
4. Complete data collected of adequate proportion of study sample?
5. If applicable, appropriate method used for data imputation?
 | **High bias**: The measurement of the PF is very likely to be different for different levels of the outcome of interest**Moderate bias**: The measurement ofthe PF may be different for different levels of the outcome of interest**Low bias**: The measurement of the PF is unlikely to be different for different levels of the outcome of interest |
| **End-point definition** | 1. Definition of end-point\*
2. Valid and reliable measurement of end-point
3. Method and setting of end-point measurement
 | 1. Clear definition of end-point provided (including duration of follow-up)?\*
2. Method reliable and valid?
3. Same method used for all patients?
 | **High bias**: The measurement of the end-point is very likely be unreliable and/or not the same method was used for all patients**Moderate bias**: The measurement of the end-point may be unreliable and/or not the same method was used for all patients**Low bias**: The measurement of the end-point is (very) likely be reliable and the same method was used for all patients |
| **Study confounding** | 1. Important confounders measured
2. Definition of the confounding factor
3. Valid and reliable measurement of confounders
4. Method and setting of confounding measurement
5. Method used for missing data
6. Appropriate accounting for confounding
 | 1. Are treatment and age at diagnosis assessed as possible confounding factors?
2. Clear definition of potential confounding variables provided?
3. Method valid and reliable?
4. Same method used for all patients?
5. If applicable, appropriate method used for data imputation?
6. Accounted for potential confounders in study design or in analysis?
 | **High bias**: The observed effect of the PFon the outcome is very likely to be distorted by another factor related to PF and outcome**Moderate bias**: The observed effect of the PF on outcome may be distorted by another factor related to PF and outcome**Low bias**: The observed effect of the PF on outcome is unlikely to be distorted by another factor related to PF and outcome |

|  |  |  |  |
| --- | --- | --- | --- |
| **Domains** | **Issues for consideration** | **Prompting questions** | **Ratings** |
| **Statistical analysis and reporting** | 1. Presentation of analytical strategy
2. Model development strategy
3. Reporting results
 | 1. Are hazard/odd ratio and confidence intervals reported?
2. Strategy for model building appropriate? Or based on a conceptual framework?
3. Selected model adequate for design of study?
4. Were all the variables considered on the statistical analysis described, even if they were not included in the final model?
 | **High bias**: The reported results are verylikely to be spurious or biased related to analysis or reporting**Moderate bias**: The reported results may be spurious or biased related to analysis or reporting**Low bias**: The reported results are unlikely to be spurious or biased related to analysis or reporting |

\*: Not included in QUIPS tool, since these items were already assessed during the full text screening; PF: prognostic factor.

****

**Figure S1.** Funnel plot showing the number of DCIS patients with a subsequent ipsilateral invasive breast cancer (iIBC) event included in the study related to the effect of a specific prognostic factor for DCIS patient with subsequent iIBC compared to DCIS patients without a recurrence. This funnel plot includes all factors that were assessed for their possible association with iIBC in all studies.



**Figure S2.** continues on next page.

Continued



**Figure S2.** continues on next page.

Continued



**Figure S2.** Forest plots of the meta-analyses performed for the prognostic factors and their association with subsequent invasive breast cancer that were reported by more than 1 high quality study: **(a)** menopausal status (pre- vs. post-); **(b)** Race and/or ethnicity (Asian vs. non-Hispanic White); **(c)** Race and/or ethnicity (Hispanic vs. non-Hispanic White); **(d)** Race and/or ethnicity (African-American vs. non-Hispanic White); **(e)** detection method (palpation vs. mammography); **(f)** margin status (involved vs. clear); **(g)** histological grade (poor vs. well differentiated); **(h)** histological grade (intermediate vs. well differentiated); **(i)** calcification (present vs. absent); **(j)** necrosis (present vs. absent); **(k)** ER (positive vs. negative); **(l)** PR (positive vs. negative); **(m)** HER2 (positive vs. negative); **(n)** Ki67 (positive vs. negative); **(o)** p16 (high vs. low); **(p)** p53 (positive vs. negative); **(q)** COX-2 (positive vs. negative); **(r)** intrinsic subtype (HR+ HER2+ vs. HR+ HER2-); **(s)** intrinsic subtype (HR- HER2+ vs. HR+ HER2-); **(t)** intrinsic subtype (HR- HER2- vs. HR+ HER2-); CI, confidence interval.