RANZCR **BIG** scientific meeting Queenstown, NZ, 5-8 April 2017

Evidence on Tomosynthesis for Population Breast Screening

Professor Nehmat Houssami

SYDNEY MEDICAL SCHOOL

Sydney School of Public Health

Sydney Medical School

THE UNIVERSITY OF SYDNEY

Presentation will focus on tomosynthesis for **population** breast cancer screening

- Overview of tomosynthesis for population screening
- Not looking at reader (observer) studies based on cancer-enriched series or when used for recalled cases, nor other applications of tomosynthesis (diagnosis, staging)

Acknowledgement

- I am supported by a National Breast Cancer Foundation (NBCF Australia) ***Breast Cancer Research Leadership Fellowship***

2

Tomosynthesis for **population** breast cancer screening

RCTs: several in progress

Prospective comparative trials

- organised population-based Euro programs
- double-reading
- mostly biennial screening
- relatively low recall (at 2D)
- within woman comparisons (keeps woman & cancer characteristics 'constant')

Retrospective evaluations

- US screening practice, radiology or breast services
- single-reading
- mostly annual screening;
- relatively high recall (at 2D)
- compare 'different' groups of screened women (could introduce bias)

3

Tomosynthesis for **population** breast cancer screening
Evidence we have

Prospective non-randomised screening trials

- Oslo (Interim): Several papers 2013-15 (*interim* data, different comparisons)
- STORM (final results): Several papers 2013-14 (different analyses)
- Malmö Breast Tomosynthesis Screening Trial - (Interim report 2015)
- STORM 2 (final results): 2016
- **Consistent evidence of significantly increased BC detection: 1.7-2.7/1000** screens
- Heterogeneous data for recall or *false-recall* : modest reductions* or slight increase (overall recall applied in trial), or slight increase from 3D-only, or report a 'mix' (pre or post arbitration meetings)

*reductions calculated conditional to 3D-mammog-positive, or false-positive scores pre-arbitration …extent it will translate in standard organised screening n*ot so clear*

4

Prospective trials Evidence tables modified & updated from Chapter by Zackrisson & Houssami (Breast Cancer Screening, Elsevier 2016

			Incremental CDR		False recall
Ciatto et al 2013 [STORM] Lancet Onc & **Houssami** et al 2014 [STORM follow-up] EJC	Prospective trial (7,292) population-based, Italy, compares 2D and 2D/3D screening (paired data); sequential double-reading, recall by either reader **Extended analysis of STORM** includes first year follow-up for interval cancers	2D: 5.3 2D/3D: 8.1 P<0.001 2D double-read: 5.3 vs 2D/3D single-read 7.5 P<0.001	↑ **2.7/1000** ↑ **2.2/1000**	Recall for 2D-alone or 2D/3D: 5.5% Recall conditional to 2D/3D-positive: 3.5% (analytic estimate 17% decrease in FP recalls)	↓ **2.0%*** **In trial increased *overall* recall by ~ 1%**
Skaane et al, 2013 [Interim Oslo trial data] Radiology **Skaane** , 2013 Euro Radiol	Prospective trial (12,631) population- based, Norway, comparing 2D and 2D/3D screening (paired data: double-reads with arbitration meeting). **Analysis of double-reading Oslo trial**	2D: 6.1 2D/3D: 8.0 (reader-adjusted) P =0.001 2D: 7.1 2D/3D: 9.4 P<0.001	↑ **1.9/1000** ↑ **2.3/1000**	2D: 6.1% 2D/3D: 5.3% (15% decrease of FP, adjusted for reader) (pre-arbitration FP scores)	↓ **0.8%** ↓ 1.8% of FP scores but **increased *overall* recall rate by 0.8%**
Lång et al 2015 [Interim MTST data], Euro Radiol	Prospective trial (7,500 from target 15,000) random sample invited in population-based program, Sweden: 2D-mammog (2 views) versus **stand-alone 1-view 3D-mammog (DBT).** Independent double-reads	2D (2-view): 6.3 3D (1-view): 8.9 P<0.0001	↑ **2.6/1000**	Overall recall rate (after arbitration): 2D: 2.6% 3D: 3.8% P<0.0001	↑ **0.9%**
Bernardi et al 2016 [STORM 2] Lancet Onc	Prospective trial (9,672) population-based, Italy, compares 2D with 2D/3D screening or 2Dsynthetic/3D; sequential double-reading X 2 parallel arms (paired data x2) , recall at any read	2D: 6.3 2D/3D: 8.5 P<0.001 2D: 6.3 2Dsynthetic/3D: 8.8 P<0.001	↑ **2.2/1000** ↑ **2.5 /1000**	2D: 3.42% 2D/3D: 3.97% P<0.001 2D: 3.42% 2Dsynthetic/3D: 4.45% P<0.001	↑ **0.55%** ↑ **1.0%**

5

Tomosynthesis for **population** breast cancer screening: Evidence we have

Retrospective studies (many & increasing) – 10 studies at March 2017

- Retrospective studies give slightly different evidence: how much due to screening setting/practice differences? How much due to different study methods (less direct comparisons):
- relatively less effect on **incremental CDR** 0-2.1/1000 (one showed non-sig reduction in CDR), median 1.2/1000
- Much more consistent and **marked reduction in recalls** (absolute range 1.4-7.3%) where higher FPR at 2D-Mam screening → stronger argument of 'harm reduction' from 3D in USA screening setting

6

Tomosynthesis for **population** breast cancer screening: Evidence we do NOT have

***Key evidence gaps*: Efficacy & Outcomes** beyond initial detection measures

- Mortality reduction (incremental)→ **most relevant measure of screening benefit** but (?feasible) limited ability/power to measure modest increments) & unlikely decisions can wait 10+ years: indirect methods (surrogate measures; modelling)
- Impact on screening outcomes (surrogates for mortality benefit) more feasible and SHOULD be a priority for new research:
  - Cancer characteristics (predicting incremental mortality benefit from ca stage distributions & ca features – but requires some assumptions)
  - Interval cancer rates (compare using valid methods!)
  - Advanced cancer rates (compare using valid methods) – absolute not proportions

7

Tomosynthesis for **population** breast cancer screening : Evidence we do NOT have

Key evidence gaps: Screen-detection measures

- What is **effect of DBT at repeat screening**? Most of the evidence from 1st ever tomo screen (even if repeat screen for woman): incremental ca detection maybe much less at repeat tomo screen - 'prevalence effect':
- Early evidence suggests increased BC detection maintained at repeat screen (*example*: STORM-2); one study showed sustained reductions in recall at repeat DBT screens vs a historical comparison group (McDonald, JamaOncology 2016)
- Subgroups: younger women, (moderately) increased risk – limited knowledge from subgroup analyses; one retrospective study shows better incremental ca dx in <50 years than in > 50 (McCarthy et al JNCI 2014)
- Program **sensitivity**: Initial ca detection + interval Ca data→ **screening program sensitivity** (initial ca detection = comparative sensitivity at screening)

8

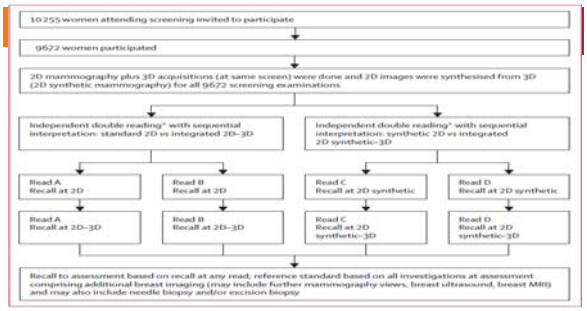
STORM 2, (Bernardi et al) *Lancet Oncology* June 2016

STORM 2 recruited in Italy

Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study

Daniela Bernardi, Petra Macaskill, Marco Pellegrini, Marvi Valentini, Carmine Fantò, Livio Ostillio, Paolina Tuttobene, Andrea Luparia, Nehmat Houssami

9



All had 2D + 3D mammography acquisitions (& synthetic 2D) operated in COMBO© mode with C-View™)

10

STORM 2: BC detection (Bernardi et al, *Lancet Oncology* 2016) All screened women

	Detected cancers, n	CDR per 1000 screens (95% CI)	p value*	Incremental CDR per 1000 screens (95% CI) attributed to integrating 3D screening vs 2D alone†
All screening participants (n=9672, analysed as n=9677)‡				
Standard digital 2D mammography	61	6·3 (4·8–8·1)	–	–
Integrated 2D-3D mammography	82	8·5 (6·7–10·5)	<0·0001	2·2 (1·2–3·3)
Integrated 2D synthetic-3D mammography	85	8·8 (7·0–10·8)	<0·0001	2·5 (1·4–3·8)

2D/3D vs 2Dsynthetic/3D: no statistical difference for CDR/incremental CDR (P=0.58)

CDR in subset of 1771 (18%) **repeat 2D/3D** screens (participated in STORM 1 & 2)
- 2D alone (12 cancers): 6.8/1000
- 2D/3D (15 cancers), 2Dsynthetic/3D (15 cancers): 8.5/1000 (+1.7)

11

Breast screening with tomosynthesis: Evidence we do NOT have

Key evidence gaps

- **Issues that matter to real-life programmatic screening & policy**:
  - Cost-effectiveness studies (high quality, not industry-funded, not assessment data)
  - Logistics & infrastructure for broader adoption in population screening
  - Impact on **screen-reading time**: HUGE problem for screening programs → Explore alternate approaches to screening delivery *Re-think mammography screening package*:
    - ✓ Examples: MBTST (3D-alone/ Less views ?)
    - ✓ modify screen-reading strategy: STORM & STORM 2 data for 2D (double-read) vs 2D/3D (single read) (Houssami EJC 2014; Houssami Cancer Epi 2017)

12

Rethink screening model using 3D

STORM 2, secondary analysis

N. Houssami et al./Cancer Epidemiology 47 (2017) 94–99

Table 2
Breast cancer detection rates and incremental cancer detection for single-reading of 3D-mammography versus double-reading of standard 2D-mammography in the STO trial.

Single-reading integrating 3D-mammography versus double-reading of standard 2D-mammography (n = 9672 analyzed as 9677)[a]	No. of detected cancers	Cancer detection rate (CDR) per 1000 screens (95% CI)	P[b]	Incremental CDR/1000 attributed to integrating 3D screening (95% CI)[c] versus 2D alone
Standard double-reading of 2D-mammography	61	6.3 (4.8, 8.1)	–	–
Single-reading of 3D with 2D mammography	79	8.2 (6.5, 10.2)	P < 0.001	1.9 (0.9, 2.9)
Single-reading of 3D with 2Dsynthetic mammography	81	8.4 (6.7, 10.4)	P < 0.001	2.1 (1.0, 3.3)

STORM , secondary analysis, from Houssami et al European Journal of Cancer 2014

Secondary outcome: Comparison of 2D/3D-mammography with single-reading and 2D-mammography with double-reading

2D-mammography, double-reading	39	5.3 (3.8, 7.3)	
Integrated 2D/3D-mammography, single-reading	55	7.5 (5.7, 9.8)	2.2 (1.0, 3.6)

[†] *p*-Value for Mc Nemar's test for paired binary data. P<0.001

13

Breast Cancer *Screening:Viewpoint of the IARC Working Group; NEJM June 2015*

- There is *sufficient evidence* that 2D-mammography + tomosynthesis increases BC detection rates compared to 2D-mammography alone
- There is *limited evidence* that 2D-mammography + tomosynthesis reduces false-positive recall compared to 2D-mammography alone
- There is ***inadequate*** *evidence* that 2D-mammography + tomosynthesis (compared to 2D-mammography alone):
  - reduces the rate of interval cancers
  - reduces breast cancer mortality
- E*vidence* that 2D-mammography + tomosynthesis (from dual acquisition) increases radiation dose compared to 2D-alone

14

Population breast screening: tomosynthesis-based screening for future

Additional (high-quality) research needed, ***RCTs would be valuable****!!!*

Evidence: screen-detection measures	Evidence: optimal 3D-screening model	Evidence: Health economics	Evidence: impact on longer-term outcomes or screening benefit
▪ Available & increasing ▪ Heterogeneous but definitely favours 3D … limited data for repeat screening. ***2D/3D better at screen-detection than 2D alone***	▪ Limited but likely to increase next few years ▪ Critical to re-think current mammography screening models	Limited	▪ Not available ▪ Requires larger data-sets ▪ Requires collaboration ▪ Requires long follow-up ▪ May require RCTs

15

Thank you for inviting me

nehmat.houssami@sydney.edu.au

16