

HIPAA-Safe Clarity SQL Checklist

Epic Clarity | techfitflow.com | Not legal advice. Your privacy officer decides what is permitted.

1. Pre-run checklist

Check	Question	Pass condition
Purpose	Can I state the decision this supports?	One sentence in the ticket
Access	Do I have approved, role-based access?	Request on file
Minimum	Is the SELECT list limited to needed columns?	No name, MRN, CSN, full dates without reason
Aggregate	Can counts answer it?	Aggregate unless justified
Small cells	Are cells under the threshold suppressed?	HAVING clause applied; check differencing
Destination	Where do results live?	Secured share; never email or personal device
Retention	When do I delete them?	Date in the ticket

2. Identifier map (standard Clarity names; verify in your dictionary)

HIPAA identifier	Typical Clarity location	Safe handling
Name	PATIENT.PAT_NAME	Drop
MRN	PATIENT.PAT_MRN_ID	Drop or random surrogate key
Dates	BIRTH_DATE, CONTACT_DATE, admit/discharge	Year only (Safe Harbor)
Geography	ADD_LINE_1, CITY, ZIP	ZIP3 only if area > 20,000; else 000
Phone/email	HOME_PHONE, EMAIL_ADDRESS	Drop
Encounter/account no.	PAT_ENC_CSN_ID, HSP_ACCOUNT_ID	Drop from exports
Age over 89	Computed from BIRTH_DATE	Roll up to 90+

3. Three paths to patient data in analytics

Path	Removes	Keeps	Approval
PHI for operations	Nothing	All, minimum necessary	Role access + purpose
Limited data set	Direct identifiers	Dates, city, state, ZIP	Data use agreement
Safe Harbor	All 18 identifiers	Year, state, ZIP3 (rule)	None once truly de-identified
Expert Determination	Per expert	More detail	Written expert determination

4. Do not

- ☐ Hash an MRN and call the data de-identified. Use random surrogate keys.
- ☐ Email row-level PHI, or attach it to Jira.
- ☐ Paste results into chat tools or AI assistants that are not approved.
- ☐ Use "anonymous" extracts that still contain dates and ZIP codes.
- ☐ Assume a test environment is scrubbed. Confirm it.

5. SQL patterns

Aggregate with small-cell suppression

```
SELECT d.DEPARTMENT_NAME, COUNT(DISTINCT e.PAT_ID) AS patients
FROM PAT_ENC e JOIN CLARITY_DEP d ON d.DEPARTMENT_ID = e.DEPARTMENT_ID
WHERE e.CONTACT_DATE >= @start AND e.CONTACT_DATE < @end
GROUP BY d.DEPARTMENT_NAME
HAVING COUNT(DISTINCT e.PAT_ID) >= 11;  -- use your policy threshold
```

Random surrogate key (restricted schema)

```
INSERT INTO rpt_secure.PAT_SURROGATE (PAT_ID, SURROGATE_KEY)
SELECT p.PAT_ID, NEWID() FROM PATIENT p
WHERE NOT EXISTS (SELECT 1 FROM rpt_secure.PAT_SURROGATE s WHERE s.PAT_ID = p.PAT_ID);
```

Masked view and grants

```
CREATE VIEW rpt_safe.vw_encounter_deid AS
SELECT s.SURROGATE_KEY AS person_key, YEAR(e.CONTACT_DATE) AS contact_year, e.DEPARTMENT_ID
FROM PAT_ENC e JOIN rpt_secure.PAT_SURROGATE s ON s.PAT_ID = e.PAT_ID;
GRANT SELECT ON SCHEMA::rpt_safe TO analyst_role;
DENY  SELECT ON PATIENT TO analyst_role;
```

6. Data access request template

Field	Your entry
Requester and role	
Purpose / decision supported	
Legal basis (operations, QI, research)	
Data needed (fields, range, population)	
Identifiers needed and reason	None / list
Granularity (aggregate or row-level, why)	
Environment and result location	
Recipients	
Retention / delete date	
Approver and date	

7. If it goes wrong

Stop sharing. Record who, what, when, where. Report to the privacy office at once. Preserve evidence as instructed. Let the privacy officer decide whether it is a breach.

Sources: HHS De-identification guidance; HHS Minimum Necessary Requirement; 45 CFR Parts 160 and 164; Epic Clarity Data Dictionary (UserWeb).