

Figure S1 Example of data services designer (version 14.2).

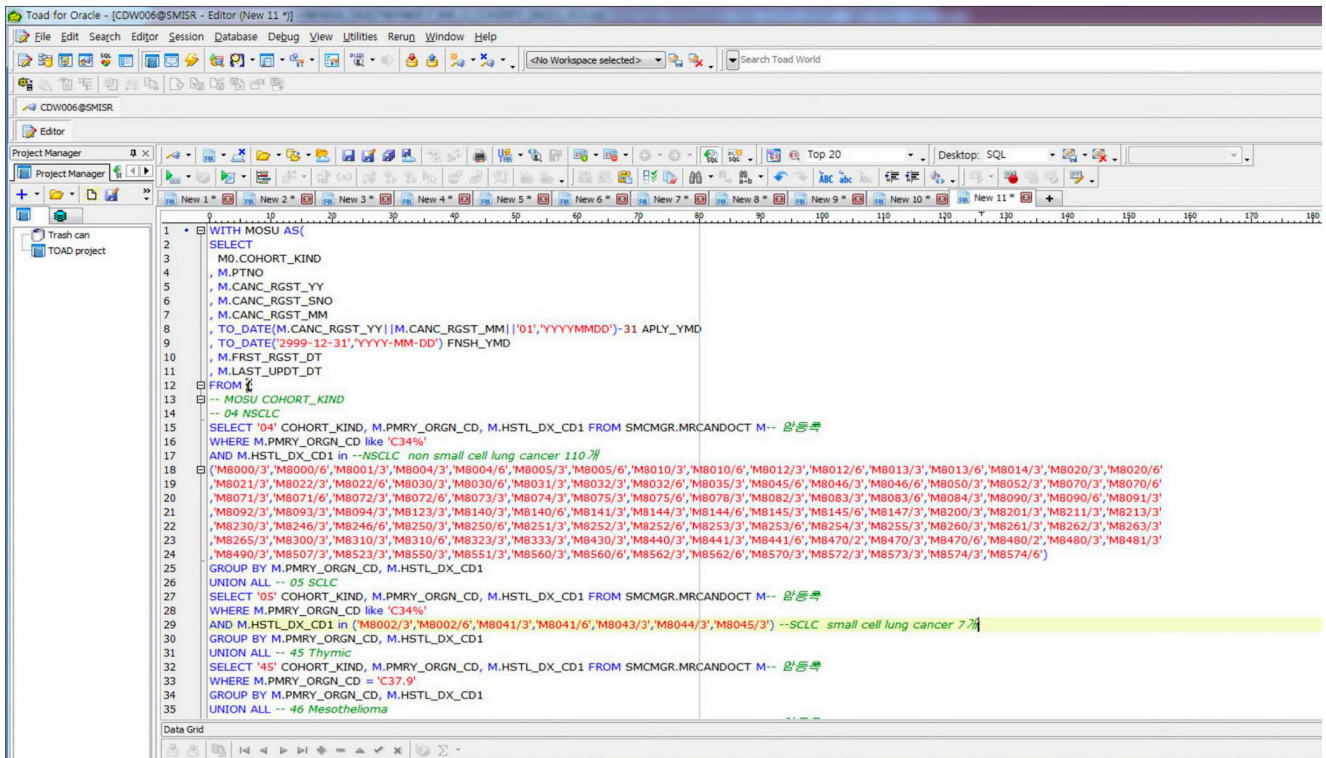


Figure S2 Example of algorithm for data extraction.

Table S1 Median OS according to clinical stage in NSCLC

| | Events/N | Median OS (95% CI) | 24 months | 60 months |
|------|------------|--------------------|-----------|-----------|
| IA | 718/5,148 | NR | 95% | 84% |
| IB | 573/2,201 | 94.1 (82.8–105.4) | 86% | 66% |
| IIA | 542/1,303 | 61.1 (55.0–67.2) | 74% | 51% |
| IIB | 375/803 | 51.4 (43.3–60.5) | 65% | 45% |
| IIIA | 1579/2,546 | 27.0 (24.9–29.1) | 54% | 34% |
| IIIB | 1176/1,575 | 15.9 (14.6–17.2) | 38% | 15% |
| IV | 6363/8,506 | 13.4 (12.9–13.8) | 33% | 10% |

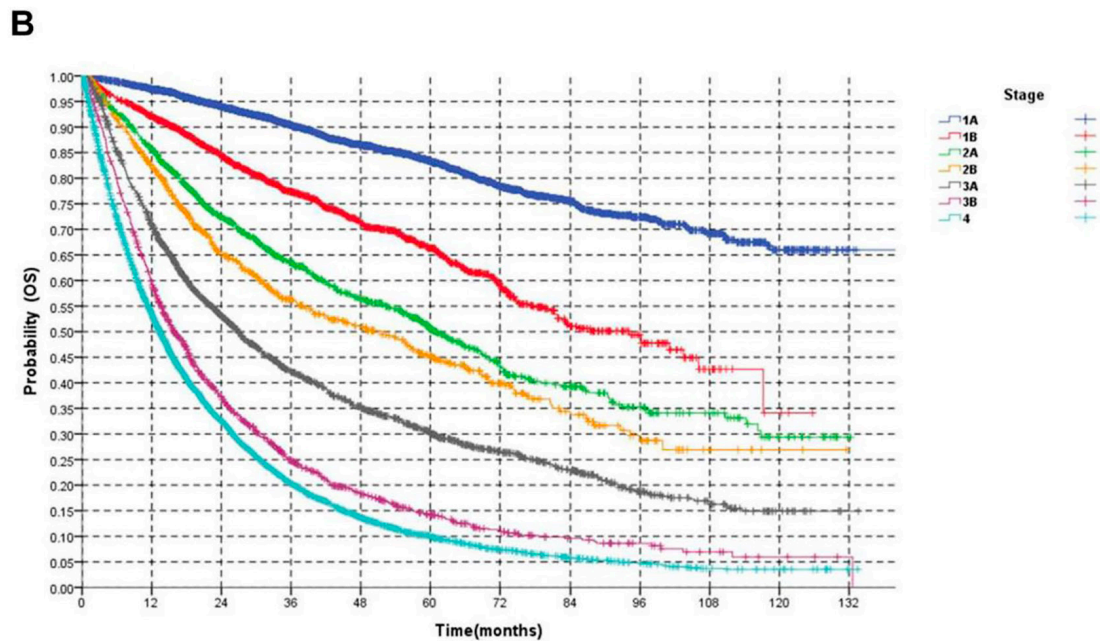
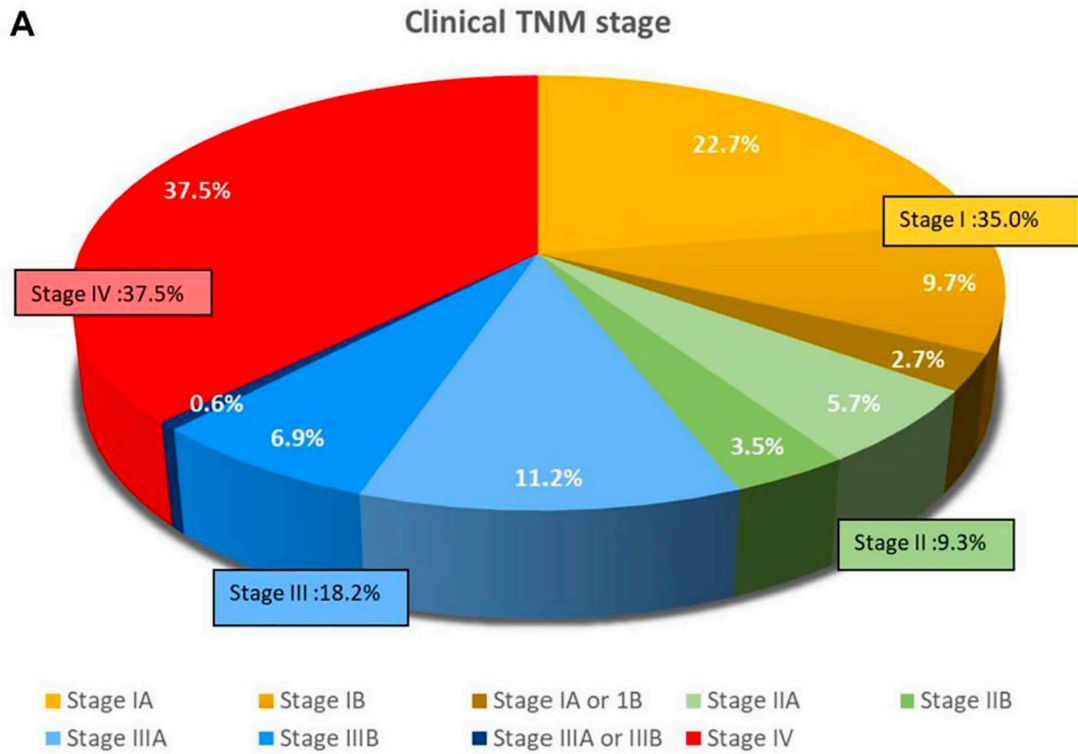


Figure S3 Examples of analysis in ROOT project. (A) Clinical TNM staging system for NSCLC (n=22,719). (B) Survival curves for overall survival in NSCLC patients. NSCLC, non-small cell lung cancer.