Adding Up the Shares of First Cancer Treatment by Type Gives 132.18% for All Sites — Because One Person Receives More Than One Treatment; Counted Again by Combination, It Sums to 100. For Breast, "Surgery" Is 78.00% but "Surgery Only" Is 9.77%, and Surgery/Endoscopy Combined with Drugs Is 42.50%. The Denominator of 972,956 Is the 993,469 Cases Minus 20,513 Known Only from Death Information [M100]

Share by type ≠ share receiving only that treatment ── What separates them is whether each person is counted once, or once for each type of treatment

Yuuki YAMAGISHI (山岸 勇輝) ／ Independent Researcher
ORCID 0009-0006-8587-3450
September 2026 ／ CC BY 4.0

Abstract

Using the 2023 National Cancer Registry's first-treatment tables (Table 6-A: share by type of treatment; Supplementary Table a3-A: share by combination of treatments; both by site), the incidence and data-quality tables, and the registry's overview and glossary, this paper looked at what the shares of first treatment measure. No new mathematical theorem and no new law are claimed. The numbers here summarize the national population; they are not any one person's outlook and do not show whether a treatment is good or bad.

First, **the denominator of the first-treatment shares (number tabulated) is 972,956 for all sites, the 993,469 cases minus those known only from death information (DCO, 2.0648% = 20,513).** For all 27 site rows, cases × (1 − DCO%) matched the number tabulated within 1 person (Section 3).

Second, this is the core of the paper. **Adding up the shares by type in Table 6-A (surgical, laparoscopic, and endoscopic treatment; radiation, chemotherapy, and endocrine therapy; other;**

no treatment/unknown) gives 132.18% for all sites, and more than 100 in all 27 rows. When one person receives several types of treatment, they are counted in each type. The mean number of types recorded for the 78.07% with treatment is 1.412, ranging by site row from 1.050 (skin) to 2.372 (brain and central nervous system) (Section 4).

Third, **Supplementary Table a3-A counts the same people once each by combination, and its 13 combinations sum to 100 in all 27 rows. For breast, "surgery" in Table 6-A is 78.00%, but "surgery only" in the supplementary table is 9.77%, "surgery/endoscopy + drugs" is 42.50%, and "surgery/endoscopy + radiation + drugs" is 22.81%.** "X% received surgery" is different from "X% received only surgery" (Section 5).

Fourth, **the two tables are different ways of counting the same people, and in all 27 rows "sum of combinations including that type ≤ share by type ≤ that plus other combinations" holds (radiation for all sites: 11.56 ≤ 11.67 ≤ 12.58). "No treatment/unknown" has the same value in both tables (21.93% for all sites)** (Section 3).

Fifth, the treatment items on the registry report form are coded "performed at this facility / not performed at this facility / unknown whether performed," and "no treatment/unknown" in the tables combines those who had no treatment and those for whom it is unknown. The tables cannot separate the two (Section 6).

Sixth, what separates them is **whether each person is counted once, or once for each type of treatment**. When reading shares of first treatment, first check whether they are shares by type (summing to more than 100) or shares by combination (summing to 100).

Scope note: **This paper claims no new mathematical theorem and no new law** — shares of first treatment, cases, and DCO are in public tables, and this paper only added them up and checked the bounds between combinations and types and how the denominator is formed. **This paper is a structural interpretation of the shape of numbers in public aggregate data; it is not medical advice, not a diagnosis, and not a treatment recommendation. The author is not a physician. Decisions about individual health should be discussed with a physician or a medical institution. This is a preprint that has not been peer reviewed. Mindful that people currently in treatment and their families may read it, this paper states first that its numbers summarize the national population and are not any one person's outlook. It does not discuss whether treatments work or are good or bad, does not compare or recommend treatments, and does not compare hospitals. Figures by site are shown only as examples of the difference between ways of counting.** "No treatment/unknown" cannot separate those with no treatment from those for whom it is unknown. The mean number of types is the mean number of recorded types; how many times the same type was given is not counted. Age, sex, and stage are not separated. **About the data**: the National Cancer Registry (Ministry of Health, Labour and Welfare) tables and documentation were taken from e-Stat and used under the Government of Japan Standard Terms of Use (Version 2.0); the sums of shares, share with treatment, mean number of types, sums of combinations, and cases × (1 − DCO%) in this paper were processed from them. What this paper adds: **lining up the two first-treatment tables (by type and by combination) and showing, in numbers for all 27 rows, that shares by type sum to more than 100, the bounds between the two tables, and that the denominator is cases minus DCO**.

1. Introduction, notation and question

The numbers here summarize the national population. They do not represent any one person's outlook or whether a treatment works.

- **Number tabulated**: the number of people in the denominator of the first-treatment shares.
- **Share by type (Table 6-A)**: the share who received that type of treatment. A person who receives several types is counted in each.
- **Share by combination (Supplementary Table a3-A)**: the share counting each person once by the combination of treatments received.
- **Share with treatment**: 100 − "no treatment/unknown." **Mean number of types**: sum of the shares of the 7 types ÷ share with treatment.
- **DCO**: cases known only from death information.

Question: What is the "share of first treatment" a share of, and what does it count?

Observation 1.1: M031 showed that the "most common cancer" changes with how sites are grouped. This paper uses the first-treatment tables of the same National Cancer Registry to see how the reading of shares changes when the same people are counted two ways, "by type" and "by combination."

---

2. Preview and design

Written honestly: the design was decided after looking at the tables.

	Content
Tables	National Cancer Registry 2023, Table 6-A (first-treatment shares), Supplementary Table a3-A (detailed first-treatment shares), Table 8-A (data-quality indicators), Table 2-A (cases by age group), and the overview, glossary, and list of tables and sites
Noticed in the preview	Table 6-A's shares sum to more than 100 by row, while Supplementary Table a3-A sums to 100. They are two ways of counting the same people
Noticed in the preview (2)	The number tabulated (972,956 for all sites) is smaller than cases (993,469). I checked whether the difference matches the DCO share
Noticed in the preview (3)	The first-treatment tables on e-Stat are by site only, not by age. This paper does not go into age
Comparison	How the denominator is formed, sums of shares by type, shares recounted by combination, and bounds between the two tables

---

3. Data and checks

Calculation 3.1 The 7 retrieved files were fixed by SHA-256, and the following were checked.

- Table 6-A and Supplementary Table a3-A have the same 27 rows (all sites and 26 sites) and the same number tabulated. "No treatment/unknown" has the same value in both. The 13 combinations of Supplementary Table a3-A sum to 100 (within 0.01) in all 27 rows.
- Number tabulated = cases × (1 − DCO%) matches within 1 person in all 27 rows (for breast (women only), women's cases and women's DCO). For all sites, cases 993,469 and DCO 2.0648% give cases × DCO = 20,513.0, against a number tabulated of 972,956 (difference 20,513). The glossary says "DCO% = cases known only from death information," an indicator of the quality of incidence statistics.
- Bounds between the two tables: in all 27 rows, sum of combinations including radiation ≤ radiation therapy ≤ that plus other combinations holds, and the same direction holds for drugs and for surgery/endoscopy (within 0.02). For radiation at all sites, 11.56 ≤ 11.67 ≤ 12.58.

4. The core — shares by type sum to more than 100

Calculation 4.1 Shares by type in Table 6-A (%)

Type	All sites	Stomach	Breast
Surgical treatment	23.50	10.73	78.00
Laparoscopic treatment	22.27	21.82	0.34
Endoscopic treatment	9.09	37.70	0.04
Radiation therapy	11.67	0.92	27.83
Chemotherapy	29.87	19.48	35.26
Endocrine therapy	11.92	0.08	56.08
Sum (incl. no treatment/unknown)	132.18	113.34	206.55

Observation 4.2: **Adding up the shares by type gives 132.18% for all sites, 113.34% for stomach, and 206.55% for breast, and more than 100 in all 27 rows.** For all sites, the 7 treatment types sum to 110.25 and "no treatment/unknown" is 21.93, giving 132.18. The mean number of types recorded for the 78.07% with treatment is 1.412, ranging by site row from 1.050 (skin) to 2.372 (brain and central nervous system) (breast 2.167). I first expected breast to be the largest, but the measurement showed brain and central nervous system.

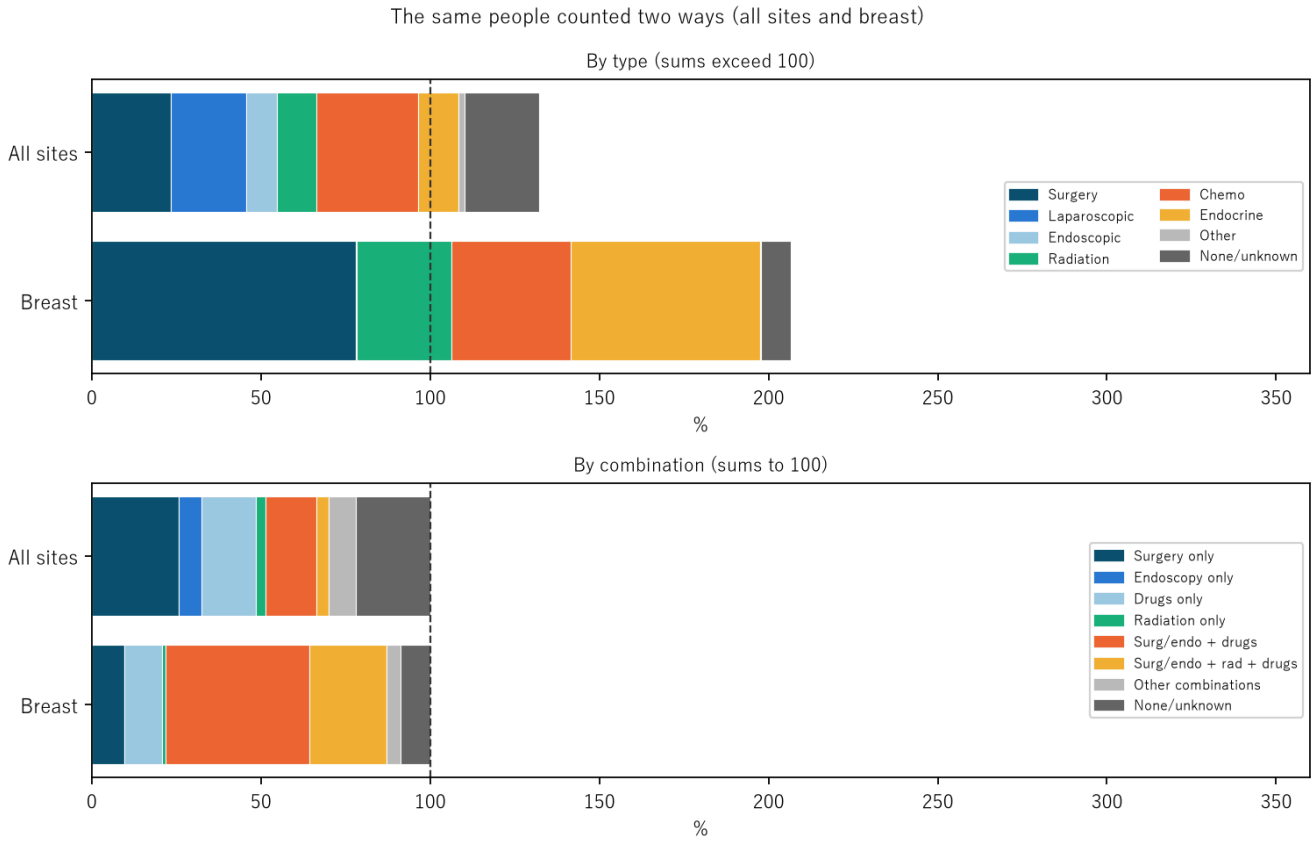


Figure 1: All sites and breast. Top, shares by type in Table 6-A stacked (summing to more than 100); bottom, shares by combination in Supplementary Table a3-A (summing to 100). The dashed line is 100%. The same color means different things in the top and bottom panels.

5. Recounting by combination

Calculation 5.1 Shares by combination in Supplementary Table a3-A (%, excerpt)

Combination	All sites	Stomach	Breast
Surgery only	25.63	20.04	9.77
Endoscopy only	6.77	35.02	0.00
Surgery + endoscopy	0.88	2.36	0.01
Drugs only	16.15	9.04	10.96
Surgery/endoscopy + drugs	15.15	10.02	42.50
Surgery/endoscopy + radiation + drugs	3.46	0.03	22.81
No treatment/unknown	21.93	22.04	8.69

Observation 5.2: **For breast, "surgical treatment" in Table 6-A is 78.00%, but counted by combination, "surgery only" is 9.77%, "surgery/endoscopy + drugs" 42.50%, "surgery/endoscopy + radiation + drugs" 22.81%, and "surgery/endoscopy + radiation" 3.09%. For stomach, against "endoscopic treatment" of 37.70% in Table 6-A, "endoscopy only" is 35.02% and "surgery + endoscopy" 2.36%.** Adding the single-type combinations (surgery only, endoscopy only, radiation only, drugs only) gives 51.36% for all sites (0.658 of those with treatment) and 21.80% for breast (0.239). For all sites, "surgery only" at 25.63% is larger than "surgical treatment" in Table 6-A at 23.50%. The tables do not say whether "surgery" in the supplementary table refers only to surgical treatment, and this paper's bounds check (Section 3) combined surgery and endoscopy (any of surgical, laparoscopic, or endoscopic).

---

6. "No treatment/unknown" and "performed at this facility"

Observation 6.1: According to the registry overview, the treatment items on the report form — surgical, laparoscopic, and endoscopic treatment, radiation, chemotherapy, endocrine therapy, and other treatment — are each coded "1 performed at this facility / 2 not performed at this facility / 9 unknown whether performed." **"No treatment/unknown" in the tables (21.93% for all sites) combines those who had no treatment and those for whom it is unknown whether treatment was performed; the tables**

cannot separate the two. Differences in "no treatment/unknown" between sites cannot be read as the need for treatment or whether it is good or bad.

---

7. Limitations

1. **The numbers in this paper summarize the national population and are not any one person's outlook.** Whether treatments work or are good or bad is not discussed, treatments are not compared or recommended, and hospitals are not compared.
2. "No treatment/unknown" cannot separate those with no treatment from those for whom it is unknown.
3. The mean number of types is the mean number of recorded types; the number of times the same type was given is not counted.
4. The first-treatment tables on e-Stat are by site only; age, sex, and stage are not separated.
5. Cases known only from death information (DCO) are excluded from the tabulation.

Conclusion

Adding up the shares of first cancer treatment by type gave 132.18% for all sites. This is because a person who receives several treatments is counted in each type; counted again once each by combination, the same people sum to 100. For breast, of "surgical treatment" at 78.00%, "surgery only" was 9.77%. The denominator of 972,956 is the 993,469 cases minus the 20,513 known only from death information.

What separates them is one thing — whether each person is counted once, or once for each type of treatment.

When you hear that "X% received a certain treatment," **check together whether it is a share by type or by combination, and who is in the denominator.**

Data and reproduction

Data	National Cancer Registry 2023, Tables 6-A, 8-A, and 2-A, Supplementary Table a3-A, and overview and documentation e1, e2, and e3 (Ministry of Health, Labour and Welfare, e-Stat) (retrieved 26 September 2026)
Terms of use	e-Stat follows the Government of Japan Standard Terms of Use (Version 2.0). The sums of shares, share with treatment, mean number of types, sums of combinations, and cases × (1 − DCO%) in this paper were processed from it and are not numbers produced by the government
Coverage	Japan, diagnoses in 2023, all sites and 26 sites (27 rows), both sexes
Verification	`検算/M100.py` (`--zu` also draws the figure). Every number in the text is taken from `検算/M100_output.txt`
Raw data	`data/` holds the retrieved files and `SOURCES.md` (URLs, retrieval date, SHA-256, terms of use, definitions and how the calculations were set up)
Figures	`図/M100_zu1` (png, svg, English version)

References

- Ministry of Health, Labour and Welfare. National Cancer Registry, incidence counts and rates 2023, national, Tables 6-A, 8-A, and 2-A; national supplementary Table a3-A; overview and documentation e1, e2, and e3 (in Japanese). e-Stat (accessed 26 September 2026). https://www.e-stat.go.jp/
- Yamagishi, Y. (2026). Which Cancer Is "Most Common" Changes with How Sites Are Grouped — For Both Sexes, the Top Site by Cases Is Colorectum When Grouped, Lung When Colon and Rectum Are Split, and Prostate at ICD Four Characters [M031]. Zenodo. https://doi.org/10.5281/zenodo.22972431