Analysis of Time and Effort Required to Collect Data for 2004 Collaborative Stage Site-Specific Factors

Emory University
Rollins School of Public Health
Hye Mi Kim, MPH
Michael Goodman, MD, MPH
Kevin Ward, PHD, MPH

Purpose of Study

- To investigate the amount of time and effort required to collect data on collaborative stage (CS) site-specific factors (SSF) in SEER

- To assess the relation between time and effort and frequency of missing data

- Sites of interest: breast, prostate, testes, colon/rectum, and lymphoma

Methods

- For each CS SSF 40 cancer registrars were asked to

  - Score data collection difficulty ranging from 1 to 5
  - Identify the main data sources
  - Estimate average time required for each variable

- Reported amounts of time and effort were then linked to the percentage of missing information in the SEER data for the period 2004-2008

Characteristics of responders

Certified Tumor Registrar, N (%)	38 (95%)
Work for an ACOS hospital, N (%)	28 (70%)
Years of Experience (median, range)	10 (29)
Annual case load of the participants' hospital >1000, N (%)	16 (40%)

Reported difficulty: Breast cancer

CS SSF	< 4 No. of Cases (%)	≥4 No. of cases (%)
SSF1 (ERA)	39 (97.5%)	1 (2.5%)
SSF2 (PRA)	39 (97.5%)	1 (2.5%)
SSF3 (Pos. ipsilateral axillary LN)	36 (90%)	4 (10%)
SSF4 (IHC of regional LN)	33 (84.6%)	6 (15.4%)
SSF5 (Molecular studies of reg. LN)	22 (56.4%)	17 (43.6%)
SSF6 (Size of invasive component)	38 (95%)	2 (5%)

Time and effort in relation to missing data: Breast cancer

CS SSF	Percent missing	Difficulty, median (range),	Percent readily available in the records	Sources of Information (%)	Estimated time in minutes, mean (SD)
SSF1 (ERA)	9.5	1.0 (4.0)	100.0	Pathology Report (92.5), Doctor's Notes(22.5), Lab Report (10.0), Other (2.5)	2.50 (2.48)
SSF2 (PRA)	10.5	1.0 (4.0)	100.0	Pathology Report (92.5), Doctor's Notes (25.0), Lab Report (7.5), Other (2.5)	2.25 (1.88)
SSF3 (Pos. ipsilateral axillary LN)	1.3	1.0 (4.0)	92.5	Pathology Report (97.5), Doctor's Notes (17.5), Lab Report (0.0), Other (2.5)	3.20 (2.88)
SSF4 (IHC of regional LN)	60.9	2.0 (4.0)	77.5	Pathology Report (84.6), Doctor's Notes (5.1), Lab Report (0.0), Other (5.1)	3.43 (3.14)
SSF5 (Molecular studies of reg. LN)	90.4	3.0 (4.0)	45.0	Pathology Report (94.8), Doctor's Notes (5.1), Lab Report (0.0), Other (2.5)	3.58 (2.87)
SSF6 (Size of invasive component)	30.2	1.0 (4.0)	95.0	Pathology Report (95.0), Doctor's Notes (20.0), Lab Report (0.0), Other (7.5)	2.77 (2.40)

Reported difficulty: Prostate cancer

CS SSF	< 4 No. of Cases (%)	≥4 No. of cases (%)
SSF1 (PSA lab value)	39 (97.5%)	1 (2.5%)
SSF2 (PSA results)	35 (87.5%)	5 (12.5%)
SSF3 (Pathologic extension)	37 (92.5%)	3 (7.5%)
SSF4 (Apex involvement)	26 (65%)	14 (35%)
SSF5 (Size of LN metastases)	38 (95%)	2 (5%)

Time and effort in relation to missing data: Prostate cancer

CS SSF	Percent missing	Difficulty, median (range),	Percent readily available in the records	Sources of Information (%)	Estimated time in minutes, mean (SD)
SSF1 (PSA lab value)	16.4	1.0 (3.0)	97.5	Pathology Report(12.5), Doctor's Notes(75.0), Lab Report(50.0), Other(7.5)	3.53 (3.67)
SSF2 (PSA results)	11.4	2.0 (4.0)	89.2	Pathology Report(10.0), Doctor's Notes(72.5), Lab Report(37.5), Other(7.5)	3.53 (3.67)
SSF3 (Pathologic extension)	0.8	1.0 (4.0)	90.0	Pathology Report(87.5), Doctor's Notes(25.0), Lab Report(0.0), Other(5.0)	3.45 (2.73)
SSF4 (Apex involvement)	32.4	3.0 (4.0)	47.5	Pathology Report(72.5), Doctor's Notes(25.0), Lab Report(0.0), Other(7.5)	5.03 (3.74)
SSF5 (Size of LN metastases)	6.65	1.0 (4.0)	95.0	Pathology Report(90.0), Doctor's Notes(37.5), Lab Report(0.0), Other(5.0)	2.73 (2.37)

Reported difficulty: Testicular cancer

CS SSF	< 4 No. of Cases (%)	≥4 No. of cases (%)
SSF1 (AFP)	30 (75%)	10 (25%)
SSF2 (hCG)	29 (72.5%)	11 (27.5%)
SSF3 (LDH)	28 (70%)	12 (30%)
SSF4 (Radical orchiectomy)	39 (97.5%)	1 (2.5%)
SSF5 (Size of LN metastases)	32 (80%)	8 (20%)

Time and effort in relation to missing data: Testicular cancer

CS SSF	Percent missing	Difficulty, median (range),	Percent readily available in the records	Sources of Information (%)	Estimated time in minutes, mean (SD)
SSF1(AFP)	24.5	2.0 (4.0)	72.5	Pathology report(7.5), Doctor's note(32.5), Lab Report(80.0), Other(5.0)	3.92 (3.60)
SSF2(hCG)	26.8	2.0 (4.0)	70.0	Pathology report(0.0), Doctor's note(40.0), Lab Report(77.5), Other(2.0)	3.89 (3.31)
SSF3(LDH)	54.6	3.0 (4.0)	55.0	Pathology report(0.0), Doctor's Note(35.0), Lab Report(75.5), Other(3.0)	4.26 (3.47)
SSF4 (Radical orchiectomy)	1.1	1.0 (4.0)	97.5	Pathology Report(72.5), Doctor's note(47.5), Lab report(2.5), Other(12.5)	2.95 (2.01)
SSF5 (Size of LN metastases)	16.7	2.0 (4.0)	72.5	Pathology Report(87.5), Doctor's note(10.0), Lab report(2.5),Other(2.0)	3.61 (2.50)

Reported difficulty:
Colorectal cancer

CS SSF	< 4 No. of Cases (%)	≥4 No. of cases (%)
SSF1 (CEA)	35 (87.5%)	5 (12.5%)
SSF2 (Clinical assessment of regional LN)	31 (77.5%)	9 (22.5%)

Time and effort in relation to missing data: Colorectal cancer

CS SSF	Percent missing	Difficulty, median (range),	Percent readily available in the records	Sources of Information (%)	Estimated time in minutes, mean (SD)
SSF1 (CEA)	51.5	2.0 (3.0)	85.0	Pathology Report (0.0), Lab Report (90.0), Doctor's Note (37.5), Others (5.0)	3.35 (2.53)
SSF2 (Clinical assessment of regional LN)	44	3.0 (4.0)	82.5	Pathology Report (12.5), Doctors' Note (45.0) Lab Report (0.0) Other (57.5)	5.39 (4.10)

Reported difficulty:
Lymphoma

CS SSF	< 4 No. of Cases (%)	≥4 No. of cases (%)
SSF1(Associated with HIV/AIDS)	27 (67.5%)	13 (32.5%)
SSF2(Systemic symptoms)	38 (95%)	2 (5%)
SSF3(IPI score)	8 (20%)	32 (80%)

Time and effort in relation to missing data: Lymphoma

CS SSF	Percent missing	Difficulty, median (range),	Percent readily available in the records	Sources of Information (%)	Estimated time in minutes, mean (SD)
SSF1(Associated with HIV/AIDS)	99.4	3.0 (4.0)	53.0	Pathology report (2.5), Lab Report (75.0), Doctor's note (17.5), Other (7.5)	6.49 (6.39)
SSF2(Systemic symptoms)	27.4	2.0 (4.0)	100.0	Pathology report (0.0), Lab Report (77.5), Doctor's note (40.0), Other (2.0)	5.77 (4.43)
SSF3(IPI score)	90	5.0 (4.0)	15.0	Pathology report (7.5), Lab report (7.5), Doctor's Note (40), Others (25.0)	7.50 (5.86)

Conclusions

1. For some CS SSF, amount of effort required for data collection and the proportion of missing data may be considered too high

2. The major obstacle to achieving data completeness is availability of information

3. The practical applications of our findings for data collection need to be explored