

조사데이터의 품질 향상을 위한 표본설계 효율화 방안 - 장애인실태조사를 중심으로¹⁾

Efficient Sample Design to Improve the Quality of Survey Data
- Focusing on the National Survey of Disabled Persons

이혜정 | 한국보건사회연구원 부연구위원

이 글은 조사데이터의 품질 향상을 위해 한국보건사회연구원에서 수행하고 있는 장애인실태조사의 표본설계 효율화 방안을 다각적으로 살펴보았다.

장애인실태조사는 조사구를 기반으로 표본추출한 가구를 조사하며, 조사 대상은 4만 가구 이상으로 꽤 많은 편에 속한다. 점점 어려워지는 조사 환경에 대응하고 조사구를 사용하는 조사에 대한 부담을 줄일 수 있는 방안을 강구하고자 하였다. 기존 방안(조사구를 활용하는 방안)과 새로운 방안(조사구와 등록장애인 DB를 병행하는 방안), 2가지 방안에 대해 장애인 출현율을 산출할 수 있는 범위 내에서 표본조사구를 기존보다 축소할 수 있는지, 표본추출틀을 인구센서스와 더불어 등록장애인 DB도 활용할 수 있는지 등의 내용을 검토하여 각 방안의 장·단점과 보완 방안을 살펴보고 향후 과제를 제안하였다.

1. 들어가며

데이터는 4차 산업혁명 시대의 핵심 자원으로 다양한 분야에서 중요하게 활용되고 있다. 데이터의 종류는 광범위하지만, 이 글에서는 설문조사로 수집된 데이터(이하 조사데이터)를 중심으로

로 살펴보고자 한다. 조사데이터는 정책 마련, 정책 효과 평가 등에서 활용되기 때문에 조사데이터의 품질관리는 매우 중요하다. 조사데이터의 품질관리는 전 단계(표본설계, 수집, 생산, 배포)에서 지속적으로 해야 한다. 이 중 표본설계에서의 품질관리, 한국보건사회연구원에서 수행하고

1) 이 글은 이혜정, 오미애, 이민경, 손창균, 박승환, 진재현, 염아림. (2020). 『전국 단위 실태조사 표본설계 효율화 방안 연구 - 장애인실태조사를 중심으로』에 수록된 내용을 바탕으로 재구성하여 작성하였다.

있는 장애인실태조사²⁾의 표본설계 효율화 방안을 다각적으로 살펴보고자 한다.

장애인실태조사는 조사구를 기반으로 표본추출한 가구를 조사하며, 조사 대상은 4만 가구 이상으로 꽤 많은 편에 속한다. 기존 대면조사 환경의 악화, 조사 비용의 증가 추세, 조사구 변동 등과 같은 이유로 점점 어려워지는 조사 환경에 대응하고자 조사구를 사용하는 조사에 대한 부담을 줄이는 방안을 마련해야 한다. 그래서 기존 방안(조사구를 활용하는 방안)과 새로운 방안(조사구와 등록장애인 DB를 병행하는 방안), 2가지 방안에 대해 다음의 내용을 중심으로 검토하고자 한다.

첫 번째, 장애인 출현율을 산출할 수 있는 범위 내에서 표본조사구를 기존보다 축소할 수 있는지에 대한 것이다. 장애인 출현율 제시는 장애인실태조사의 목적 중 하나이므로 조사 방식이 변경되더라도 산출될 수 있어야 한다. 그래서 장애인 출현율 조사를 위한 가구 및 판별 조사의 비중을 줄일 수 있는지에 대한 모의실험을 실시해 볼 것이며, 이때 조사구 축소 비율은 다양하게 설정하여 살펴볼 것이다. 두 번째, 표본추출틀을 인구센서스와 더불어 등록장애인 DB도 활용할 수 있는지에 대한 것이다. 등록장애인 DB를 활용한 리스트 조사 병행 가능성을 검토하여 등록장애인에 대한 심층 조사 확대 가능성을 가늠해 보고자 한다.

이 글은 장애인실태조사 표본설계 현황, 표본조사구 축소를 위한 모의실험, 그리고 이를 기반으로 하여 표본설계 효율화 방안으로 ‘조사구를 활용하는 방안’ 및 ‘조사구와 등록장애인 DB를 병행하는 방안’에 대해 장·단점, 보완 방안 등을 살펴본다.

2. 장애인실태조사 표본설계 현황

장애인실태조사는 첫 조사가 1980년에 시작된 이후 2005년까지는 5년 주기로 조사하였으나, 이후에는 3년 주기로 조사를 하였으며 2021년에는 11차 조사를 완료하였다.

목표모집단은 장애인으로 정의한다. 본 조사에서의 장애인은 「장애인복지법」 제32조(장애인 등록)에 따라 장애인 등록을 한 자(이하 등록장애인)로 국한하지 않고, 장애인 등록을 하지 않았으나 장애 상태에 있는 자(이하 미등록장애인)를 모두 조사 대상에 포함한다.

조사모집단(표본추출틀)을 살펴보면, 2008년과 2020년을 제외한 나머지 조사 연도에서는 모두 연도별 최신 인구주택총조사(등록센서스) 조사구를 이용하여 대상 지역을 선정하였다. 반면 2008년에는 비장애인을 조사하는 것에 대한 지적이 있어서 등록장애인만을 대상으로 조사하였고, 등록장애인 DB를 사용하여 조사 대상자를

2) 장애인실태조사의 목적은 2가지이다. 첫 번째 목적은 우리나라 장애인구를 파악하고, 우리나라 유일의 장애인 출현율을 산출하는 것이다(김성희 외, 2018, p. 33). 두 번째는 장애인에 대한 심층 조사를 실시하여 장애인의 생활 실태 및 복지 욕구를 파악하고 복지서비스 이용 등에 대한 신뢰성 있는 데이터를 생산하며, 장단기 장애인복지정책 수립과 추진을 위한 기초 자료를 제공하는 것이다(김성희 외, 2018, p. 33).

선정하였다. 2020년 역시 코로나19로 조사 여건의 어려움이 발생하여 등록장애인 DB를 사용하였다.

장애인실태조사의 표본설계는 표본에 15개 장애유형이 모두 포함되어 장애유형별 복지서비스 욕구가 적절히 반영되어야 한다(김성희 외, 2011, p. 53; 김성희 외, 2014, p. 57; 김성희 외, 2018, p. 69). 표본추출틀은 이러한 장애유형에 대한 정보가 없어서 표본설계에서는 반영할 수 없으며, 조사모집단과 목표모집단 간에도 가구와 개인으로 추정 수준이 다른 한계점도 있다(김성희 외, 2011, p. 53; 김성희 외, 2014, p. 57; 김성희 외, 2018, p. 69). 표본 배분 단계에서 이러한 점을 보완하기 위하여 제한적인 정보이지만, 장애인과 장애인 가구는 서로 연관성이 높으므로 지역별 장애인 분포를 이용하였다(김성희 외, 2011, p. 53; 김성희 외, 2014, p. 57; 김성희 외, 2018, p. 69).

층화 기준은 2017년과 2014년은 대도시, 중소도시, 농어촌의 3개 층, 2011년은 4개 층, 2005년은 6개 층, 2000년은 5개 층, 1995년은

3개 층으로 층화하였다(〈부표 1〉 참조). 이는 시도별로 장애인 출현율이 낮은 특정 장애인을 포괄하기 위해서였다. 한편 2008년에는 전국 16개 시도별 읍면동을 집락 및 층화(3,573개 읍면동)하여 전국 16개 시도별 읍면동을 1차 추출단위로 하였고, 추출된 표본 읍면동(208개)의 등록장애인을 지역별 표본 읍면동의 15개 장애유형과 6개 장애등급으로 층화한 후 2차 추출단위인 등록장애인을 추출하였다(변용찬 외, 2009, pp. 64-65).

〈표 1〉을 보면 조사구 수는 2011년 이후 층별 등록장애인 규모를 고려하여 1,000개 수준을 추출하였고, 장애출현율을 고려하여 조사구당 45 가구를 조사하였다. 2005년 이전에는 200개 이하로 나타났는데, 2005년의 경우 통합조사구를 표본조사구로 사용하였기 때문이다. 통합조사구란 표본으로 지정된 조사구가 2000년 인구주택총조사의 조사구인 경우에는, 이 조사구부터 차례로 하나씩 조사구를 추가하여 그 누적 가구가 220가구를 초과할 때까지의 3~5개의 조사구를 통합한 것이다(변용찬 외, 2006, p. 89). 그래

표 1. 조사구 수, 대상 가구수, 조사 완료 가구수 및 완료율

연도	조사구 수	대상 가구수	조사 완료 가구수	완료율
2017년	1,000개	44,161가구	36,200가구	81.9%
2014년	1,004개	48,344가구	38,560가구	79.8%
2011년	1,000개	47,458가구	38,231가구	80.6%
2005년	187개	45,285가구	40,556가구	89.6%
2000년	200개	44,128가구	39,411가구	89.3%
1995년	200개	41,283가구	39,078가구	94.6%

주: 2008년은 표본추출틀이 등록장애인 DB이므로 제외함.

자료: 이해정 외, (2020). 전국 단위 실태조사 표본설계 효율화 방안 연구 - 장애인실태조사를 중심으로. 한국보건사회연구원. p. 25. 〈표 2-3〉.

서 187개 표본조사구에 포함된 인구주택총조사 조사구는 776개 조사구로 나타났다(변용찬 외, 2006, p. 90). 2000년은 조사구당 평균 221가구였으며(변용찬 외, 2000, p. 71), 1995년은 200개의 표본조사구에 대해 전수조사(41,283가구 대상)를 실시하였다(정기원, 권선진, 계훈방, 1995, p. 66).

가구 및 판별 조사의 조사 완료 가구수는 2017년 36,200가구로 꾸준히 감소하고 있는 것을 알 수 있다. 2005년에는 40,556가구로 조사 완료 가구수가 가장 많았다. 조사 완료율을 보면 모든 조사 연도에서 80% 이상으로 조사 성공률이 높은 편이라고 볼 수 있으며 특히 1995년에 94.6%로 가장 높게 나타났다.

표본추출방법은 등록장애인 DB를 사용하여 표본추출한 2008년(층화집락추출방법)을 제외하고는 확률비례계통추출방법을 사용하였다.

그동안 조사에서의 표본설계 현황을 살펴본 결과, 2008년에 등록장애인 DB를 표본추출틀로 활용한 경우를 제외하고 나머지 조사에서는 인구주택총조사(등록센서스) 조사구를 활용하여 비슷한 표본설계 방법으로 조사를 실시하였음을 알 수 있었다.

3. 표본조사구 축소 접근 방식

가구조사의 비중을 줄일 수 있는지를 검토하기 위하여 모의실험을 실시하였다. 가구조사는 표본설계를 통해 추출된 표본조사구 내의 가구와 접촉하여 조사하는 방식이다. 2017년 조사 자료를 이용하여 표본조사구 현황을 살펴보고, 표본조사구 축소에 대한 모의실험을 실시하였다.

가. 2017년 표본조사구 현황

2017년 조사의 경우 모집단 층화는 2016년 등록장애인 수 분포를 고려하여 권역별(3개 층, 대도시/중소도시/농어촌)로 적정 표본 규모를 산정하였다. 그런 다음 전국 17개 시도별로 층화하였고, 7개 특별·광역시를 제외한 9개 도 및 세종시 지역은 동부와 읍면부로 구분하여 3차 층화를 하였다. 그리고 조사구 특성에 따라 표본조사구를 추출하였다. 즉 층화는 권역 → 시도 → 읍면으로 구분된다고 볼 수 있다.

〈표 2〉에서 보듯이 중소도시(39.8%)와 대도시(38.0%)가 표본조사구 대부분을 차지하고 있으며 농어촌은 22.2%로 나타났다.

표 2. 권역별 표본조사구 분포

권역	표본조사구	
	개	%
대도시	760	38.0
중소도시	797	39.8
농어촌	444	22.2
전체	2,001	100.0

자료: 이해정 외. (2020). 전국 단위 실태조사 표본설계 효율화 방안 연구 - 장애인실태조사를 중심으로. 한국보건사회연구원. p. 56. 〈표 3-1〉.

〈표 3〉은 2017년 표본조사구 분포로 총 2,001개의 표본조사구를 사용하였다. 〈표 1〉의 2017년에 산출한 1,000개와 다른 까닭은 2015년 등록센서스 기반 표본추출틀 조사구 중에서 병합조사구를 활용해야 하는데, 그 당시 제공 시점이 10월로 예정되어 있어서 개별조사구를 2개 병합하여 평균 60가구로 만들어 활용했기 때문이다. 동부는 서울특별시와 경기도가 19.9%로 표본조사구 비중이 가장 높았으며, 다음으로 부산광역시(9.3%), 인천광역시(8.2%), 대구광역시(5.8%), 대전광역시(5.8%) 순으로 나타났다. 이에 반해 울산광역시(1.2%)와 세종시(0.6%)는 동부 중에서 아주 낮은 편에 속하였다. 읍면부는 전라남도가 20.2%, 경상북도는 18.9%, 경상남도가 14.0%로 높은 편이고, 세종시(1.3%), 전라북도(2.5%), 제

주도(2.6%)가 낮은 편으로 나타났다.

다음은 장애인 가구를 포함하고 있는 표본조사구의 분포에 대해 전체 장애인 가구, 미등록장애인 가구로 나누어 살펴보았다. 여기서 전체 장애인은 등록장애인과 미등록장애인을 모두 포함하는 것이다. 표본조사구에서 할당된 가구를 접촉하기 전에는 가구 내 법정장애인의 포함 여부를 알 수 없다. 장애인실태조사는 법정장애인이 있는 가구의 경우 가구 및 판별 조사뿐만 아니라 심층 조사도 작성하기 때문에 표본조사구 내에 법정장애인을 포함하고 있는 가구 현황을 살펴보는 것은 의미가 있다. 더불어 미등록장애인 가구가 포함된 표본조사구 분포도 파악하였다.

표본조사구 내 전체 장애인 및 미등록장애인 규모는 〈표 4〉와 같다. 전체 장애인의 경우 경기

표 3. 시도별 표본조사구 분포

지역	동부		읍면부		전체	
	개	%	개	%	개	%
서울특별시	272	19.9	-	-	272	13.6
부산광역시	128	9.3	-	-	128	6.4
대구광역시	80	5.8	-	-	80	4.0
인천광역시	112	8.2	-	-	112	5.6
광주광역시	56	4.1	-	-	56	2.8
대전광역시	80	5.8	-	-	80	4.0
울산광역시	16	1.2	-	-	16	0.8
세종시	8	0.6	8	1.3	16	0.8
경기도	272	19.9	72	11.4	344	17.2
강원도	40	2.9	57	8.9	97	4.8
충청북도	40	2.9	56	8.8	96	4.8
충청남도	32	2.4	72	11.4	104	5.2
전라북도	56	4.1	16	2.5	72	3.6
전라남도	48	3.5	128	20.2	176	8.8
경상북도	48	3.5	120	18.9	168	8.4
경상남도	56	4.1	88	14.0	144	7.2
제주도	24	1.8	16	2.6	40	2.0
전국	1,368	100.0	633	100.0	2,001	100.0

자료: 이해정 외. (2020). 전국 단위 실태조사 표본설계 효율화 방안 연구 - 장애인실태조사를 중심으로. 한국보건사회연구원. pp. 56-57. 〈표 3-2〉.

표 4. 표본조사구 내 전체 장애인 및 미등록장애인 규모

(단위: %)

지역	전체 장애인			미등록장애인		
	동부	읍면부	전체	동부	읍면부	전체
서울특별시	17.0	-	10.1	21.3	-	12.1
부산광역시	10.3	-	6.1	32.6	-	18.5
대구광역시	5.6	-	3.3	6.7	-	3.8
인천광역시	7.7	-	4.5	4.5	-	2.5
광주광역시	3.2	-	1.9	2.2	-	1.3
대전광역시	6.2	-	3.7	3.4	-	1.9
울산광역시	1.4	-	0.8	0.0	-	0.0
세종시	0.3	1.0	0.6	0.0	0.0	0.0
경기도	16.9	9.6	14.0	7.9	2.9	5.7
강원도	2.8	8.3	5.1	5.6	8.8	7.0
충청북도	4.8	10.4	7.1	1.1	1.5	1.3
충청남도	2.5	9.9	5.5	0.0	5.9	2.5
전라북도	6.0	2.6	4.6	4.5	0.0	2.5
전라남도	4.9	23.4	12.4	3.4	30.9	15.3
경상북도	3.4	16.9	8.9	1.1	17.6	8.3
경상남도	5.2	15.0	9.1	5.6	30.9	16.6
제주도	1.8	2.9	2.3	0.0	1.5	0.6
전국	100.0	100.0	100.0	100.0	100.0	100.0

자료: 이해정 외. (2020). 전국 단위 실태조사 표본설계 효율화 방안 연구 - 장애인실태조사를 중심으로. 한국보건사회연구원. pp. 60-61. <표 3-5>.

도(14%), 전라남도(12.4%), 서울특별시(10.1%)가 높은 편이었으며, 세종시(0.6%) 및 울산광역시(0.8%)가 낮은 편으로 나타났다. 미등록장애인의 경우 부산광역시(18.5%), 경상남도(16.6%), 전라남도(15.3%)가 높은 편으로 나타나 전체 장애인과는 다른 결과를 보였다.

앞의 결과는 2017년 실태조사의 표본조사구 현황이기에 일반화할 수는 없으나 현황 파악 시 참고 자료로 활용하는 데 의미가 있다.

나. 표본조사구 축소 관련 모의실험

모의실험은 2017년 조사의 2,001개 표본조사구를 모집단으로 설정하였다. 표본조사구의 축소

비율을 5가지(10%, 20%, 30%, 40%, 50%)로 설정하여 다양한 축소 비율에 따른 표본조사구의 축소 효과를 비교하였다.

모의실험 절차는 표본조사구를 10% 축소하는 경우를 예로 설명하면 다음과 같다. 모집단 표본조사구를 전국 17개 시도별로 층화한 후 7개의 특별·광역시(서울특별시, 부산광역시, 대구광역시, 인천광역시, 광주광역시, 대전광역시, 울산광역시)를 제외한 9개의 도(경기도, 강원도, 충청북도, 충청남도, 전라북도, 전라남도, 경상북도, 경상남도, 제주도)와 세종시는 동부와 읍면부로 층화하였다. 각 해당하는 셀(cell)마다 10%를 축소한 값을 구하였고, 이 값이 추출해야 하는 표본조사구의 개수가 되는 것이다. 예를 들면 서울의 경

표 5. 시도별 동부·읍면부별 가구 분포 및 가구수 기준 최대허용오차 한계

(단위: 가구, %)

지역	모집단		표본조사구 축소 10%		표본조사구 축소 20%		표본조사구 축소 30%		표본조사구 축소 40%		표본조사구 축소 50%	
	가구수	최대허용 오차	가구수	최대허용 오차	가구수	최대허용 오차	가구수	최대허용 오차	가구수	최대허용 오차	가구수	최대허용 오차
전국	36,200	0.515	32,520	0.543	28,987	0.576	25,328	0.616	21,767	0.664	18,095	0.729
서울특별시	동부 4,918	1.397	4,430	1.472	3,946	1.560	3,438	1.671	2,948	1.805	2,459	1.976
부산광역시	동부 2,316	2.036	2,080	2.149	1,846	2.281	1,626	2.431	1,394	2.625	1,158	2.881
대구광역시	동부 1,447	2.576	1,300	2.718	1,158	2.880	1,012	3.081	867	3.329	724	3.644
인천광역시	동부 2,026	2.177	1,826	2.294	1,628	2.429	1,411	2.609	1,214	2.813	1,012	3.082
광주광역시	동부 1,013	3.079	904	3.260	814	3.434	705	3.692	614	3.954	506	4.356
대전광역시	동부 1,447	2.576	1,303	2.715	1,158	2.880	1,012	3.080	868	3.326	723	3.646
울산광역시	동부 290	5.755	254	6.153	236	6.383	198	6.962	181	7.294	145	8.141
세종시	동부 144	8.167	126	8.736	108	9.420	108	9.420	91	10.296	72	11.556
	읍면부 144	8.167	126	8.736	107	9.488	107	9.488	89	10.408	72	11.632
경기도	동부 4,941	1.394	4,451	1.469	3,961	1.557	3,452	1.668	2,962	1.801	2,471	1.972
	읍면부 1,300	2.718	1,173	2.862	1,048	3.027	904	3.260	775	3.522	648	3.851
강원도	동부 724	3.642	652	3.839	579	4.071	509	4.346	436	4.695	361	5.162
	읍면부 1,021	3.067	912	3.245	823	3.416	717	3.661	608	3.975	519	4.303
충청북도	동부 722	3.647	650	3.844	578	4.077	505	4.361	433	4.708	360	5.166
	읍면부 1,007	3.088	900	3.267	809	3.447	703	3.698	612	3.964	502	4.378
충청남도	동부 584	4.055	529	4.259	475	4.499	403	4.883	346	5.268	292	5.734
	읍면부 1,298	2.720	1,172	2.863	1,046	3.030	902	3.264	776	3.519	648	3.849
전라북도	동부 1,009	3.085	901	3.265	810	3.443	702	3.698	612	3.961	505	4.360
	읍면부 288	5.775	252	6.177	233	6.416	198	6.967	181	7.298	144	8.181
전라남도	동부 867	3.328	777	3.516	687	3.740	615	3.953	524	4.282	432	4.714
	읍면부 2,311	2.039	2,074	2.152	1,841	2.284	1,626	2.430	1,390	2.629	1,155	2.884
경상북도	동부 871	3.321	780	3.508	689	3.734	616	3.950	526	4.272	437	4.690
	읍면부 2,165	2.106	1,949	2.220	1,732	2.355	1,513	2.519	1,301	2.717	1,080	2.982
경상남도	동부 1,020	3.068	911	3.247	821	3.421	710	3.679	621	3.932	509	4.343
	읍면부 1,597	2.452	1,432	2.589	1,271	2.749	1,126	2.921	961	3.161	797	3.471
제주도	동부 436	4.693	399	4.905	344	5.282	307	5.594	254	6.150	216	6.672
	읍면부 294	5.715	258	6.101	240	6.333	203	6.883	183	7.241	148	8.068

자료: 이혜정 외. (2020). 전국 단위 실태조사 표본설계 효율화 방안 연구 - 장애인실태조사를 중심으로. 한국보건사회연구원. pp. 66-67. <표 3-8>.

우 모집단의 표본조사구는 272개(부표 2)인데 10%를 축소한다면 추출해야 할 표본조사구는 245개($272 \times 0.9 = 244.8$)이고, 세종시도 같은 방법으로 계산하면 7개($8 \times 0.9 = 7.2$)인 것이다. 10%를 축소하는 경우 표본조사구의 추출 개수는 총 1,798개(동부: 1,230개, 읍면부: 568개)이다(부표 2). 20%를 축소하는 경우 표본조사구의 추출 개수는 총 1,602개(동부: 1,095개, 읍면부: 507개)이다. 30%를 축소하는 경우 표본조사구의 추출 개수는 총 1400개(동부: 957개, 읍면부: 443개)이다. 40%를 축소하는 경우 표본조사구의 추출 개수는 총 1,203개(동부: 822개, 읍면부: 381개)이다. 50%를 축소하는 경우 표본조사구의 추출 개수는 총 1,001개(동부: 684개, 읍면부: 317개)이다.

시도별 동부·읍면부별로 해당하는 표본조사구 개수를 단순임의추출(simple random sampling) 방법으로 최종 추출하였는데 모의실험의 반복 수는 100번으로 설정하였다.

모의실험 결과의 평가지표는 단순임의추출 가정으로 95% 신뢰수준하에서 가구수 기준 최대허용오차 한계이며 식은 다음과 같다. 최대허용오차는 표본오차(sampling error)³⁾에 대해 허용 가능한 최댓값을 의미하며, 표본오차는 표준오차(standard error)와 상대표준오차(Relative Standard Error: RSE) 등으로 보여준다.

$$\text{최대허용오차} = 1.96 \times \frac{0.5}{\sqrt{\text{해당 셀의 가구수}}} \times 100$$

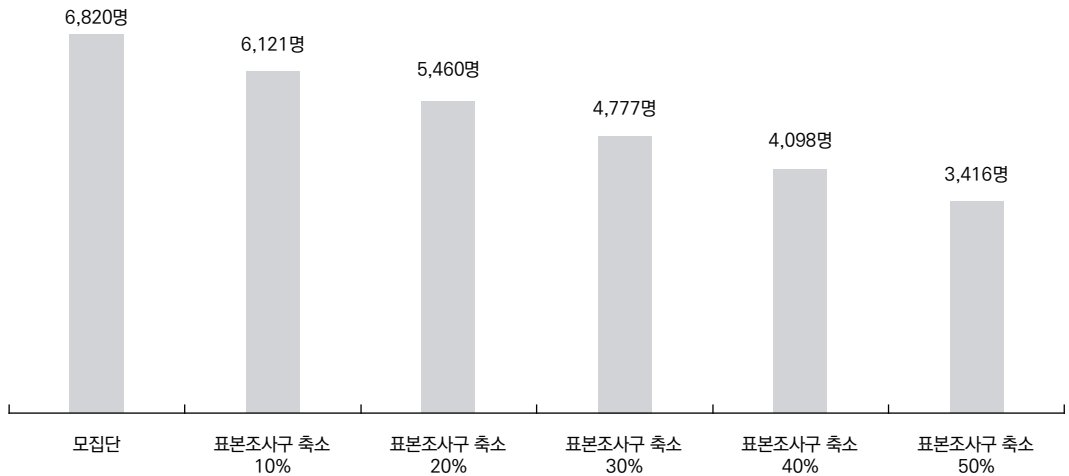
최대허용오차 한계는 모의실험이 1번 실행될 때마다 생성한 값을 누적한 후 평균값으로 제시하였다. 오차의 한계가 클수록 구간이 넓어지며 점 추정치에 대해 확신할 수 없다고 볼 수 있다. 시도별 동부·읍면부별 가구조사 대상 가구 분포 및 가구수 기준 최대허용오차 한계, 전국의 전체 장애인 규모도 함께 살펴보았다.

먼저, 시도별 동부·읍면부별 가구수 기준 최대허용오차 한계를 살펴보았다(표 5). 전국의 경우 표본조사구 축소 비율과 상관없이 모두 1% 이하로 나타났다. 표본조사구 축소 30% 이하에서는 시도별 동부·읍면부별 최대허용오차가 모두 10% 이하로 나타났다. 그러나 표본조사구 축소 40% 이상의 경우 세종시 동부 및 읍면부에서 10% 이상으로 나타났다. 그 외 나머지 지역에서는 모두 10% 이하로 나타났다.

다음은 조사구 축소 비율에 따른 전국의 전체 장애인 규모를 살펴보았다(그림 1). 전체 장애인 규모를 살펴보면 모집단은 6,820명이다. 표본조사구를 10%로 축소한 경우, 전체 장애인 규모는 6,121명으로 나타났으며 모집단 대비 89.8%를 차지하였다. 즉, 표본조사구를 10%로 축소하면 조사 대상자 규모도 약 10% 축소된다고 볼 수 있다. 표본조사구를 20%로 축소한 경우에는 5,460명(80.1%), 30%의 경우 4,777명(70.0%),

3) 표본오차란 모집단 전수조사가 아닌 표본조사를 하여 발생하는 오차이다.

그림 1. 전체 장애인 규모



자료: 이해정 외. (2020). 전국 단위 실태조사 표본설계 효율화 방안 연구 - 장애인실태조사를 중심으로. 한국보건사회연구원, p. 68. <표 3-9>를 바탕으로 재구성함.

40%의 경우 4,098명(60.1%)이고 50%의 경우 3,416명(50.1%)으로 나타났다.

이상의 모의실험 결과를 보듯이 표본조사구를 축소한 비율만큼 조사 대상 규모도 비슷한 비율로 축소되는 것을 확인하였다.

4. 장애인실태조사 표본설계 효율화 방안 검토

장애인실태조사에서 ‘조사구를 활용하는 방안’ 및 ‘조사구와 등록장애인 DB를 병행하는 방안’을 검토하였다. 첫 번째 조사구를 활용하는 방안은 현재 장애인실태조사에서 실시하고 있어서 현행 유지 방안이라고 볼 수 있다. 두 번째 조사구와 등록장애인 DB를 병행하는 방안은 기존 조사 방식에 추가로 등록장애인 DB도 활용하여 조

사하는 방안이라고 볼 수 있다. 즉, 가구 및 판별 조사의 비중은 줄이고 장애인에 대한 심층 조사를 확대하는 방안으로 볼 수 있다.

가. 조사구를 활용하는 방안(현행 방안)

조사모집단은 통계청의 인구센서스를 바탕으로 작성된 조사구로 단일추출틀이다(표 6). 일반적으로 조사구는 모집단 전수 리스트를 제외하고 통계적 추출 방법으로 확률추출이 가능함과 동시에 대표성 높은 표본추출이 가능한 표본추출틀이라는 점에서 타당하다고 볼 수 있다. 현 상황의 제한된 조사 예산 및 인프라(연구 인력, 투입 시간 등의 자원)하에서 적합한 방안으로 볼 수 있다. 표본추출은 이중추출방법(double sampling)⁴⁾

표 6. 조사구를 활용하는 방안 - 목표모집단, 조사모집단, 표본추출방법

구분	내용
목표모집단	장애인
조사모집단(표본추출틀)	등록센서스 기반 조사구(단일추출틀)
표본추출방법	이중추출방법

자료: 이해정 외. (2020). 전국 단위 실태조사 표본설계 효율화 방안 연구 - 장애인실태조사를 중심으로. 한국보건사회연구원. pp. 8-9. <요약표 1-1>.

을 이용한다. 장애인에 대한 정보가 사전에 없고 장애인 발현율이 낮으므로 1상(first phase)에서 대규모 가구 및 판별 조사를 하여 장애인을 식별한다. 2상에서는 1상에서의 조사를 통해 판별된 장애인을 대상으로 심층 조사를 실시한다. 이러한 방법은 2개의 조사를 해야 하므로 조사 비용이 많이 소요되기는 하지만 장애인을 모두 포괄할 수 있다. 다만, 1상에서의 표본 규모 관련 비용과 시간을 절약할 수 있는 최적값 산출에 대한 심층 연구가 필요하다고 생각한다.

나. 조사구와 등록장애인 DB를 병행하는 방안(새로운 방안)

조사모집단은 인구센서스를 바탕으로 작성된

조사구와 등록장애인 DB로 이중추출틀이다(표 7). 이중추출틀을 기반으로 한 표본추출방법은 조사구와 등록장애인 DB 명부 추출 시 읍면동 기준을 동일하게 하는 것이 효율적이라고 생각한다. 동일 읍면동을 사용 시 집락들을 구성하고 있는 조사 단위들의 이질적인(heterogeneous) 특성과 집락 간에 대한 동질적인(homogeneous) 특성도 확인할 필요가 있다. 조사 단위들이 이질적일수록 집락 간 변동이 줄어들어서 동일한 조사 비용으로 보다 효율적인 추정이 가능하기 때문이다.

한편, 한정된 조사 예산이므로 조사구와 등록장애인 DB에서의 표본 배분을 검토해야 한다. 그래서 모의실험을 통해 다양한 표본조사구 축소

표 7. 조사구와 등록장애인 DB를 병행하는 방안 - 목표모집단, 조사모집단, 표본추출방법

구분	내용
목표모집단	장애인
조사모집단(표본추출틀)	등록센서스 기반 조사구 및 등록장애인 DB(이중추출틀)
표본추출방법	층화다단계추출방법

자료: 이해정 외. (2020). 전국 단위 실태조사 표본설계 효율화 방안 연구 - 장애인실태조사를 중심으로. 한국보건사회연구원. pp. 8-9. <요약표 1-1>.

4) 이중추출방법은 표본추출틀에서 층화할 때 필요한 보조 변수 정보가 충분히 들어 있지 않아서 표본추출틀이 정확하지 않을 때 사용한다. 1차로 대규모의 표본을 추출하여 층화변수로 사용할 수 있으면서도 응답이 간편한 변수들에 대해 조사를 실시한 다음, 이를 근거로 1차 표본단위를 층화하여 각 층에서 일부의 2차 표본을 추출하는 방법이다(손창균, 홍기학, 이기성, 2006, p. 29).

비율에 따른 최대허용오차 한계를 살펴보고, 축소된 조사구 조사로 가용할 수 있는 조사 예산으로 추가 심층 조사가 가능한 등록장애인 규모도 가능해 보았다(‘3. 표본조사구 축소 접근 방식’ 참조). 모의실험 결과를 종합해 보면 표본조사구를 축소할 비율만큼 조사 대상 규모도 비슷한 비율로 축소되는 양상을 보였다. 가구수 기준 최대허용오차 한계 및 조사 대상 규모의 결과를 통하여 기존 표본조사구 규모의 30% 이하까지는 축소가 가능하다고 볼 수 있다. 단, 조사의 정확도 및 조사 비용을 종합적으로 고려한 최대허용오차 기준 마련 및 조사구 조사를 통해서만 가능한 미등록장애인의 구축 방안 등에 대한 심층 연구를 진행해야 한다.

현재 실시 중인 국내 장애인 관련 실태조사 예산 집행을 기반으로 하여, 2017년 장애인실태조사 예산 기준으로 심층 조사가 추가 가능한 표본수를 산출해 보았다(표 8). 단, 조사 사례비 수준에 따라 예산 변동이 가능하고, 조사구와 등록장애인 DB에서 추출된 표본의 중복 가능성을 반영

해야 하는 등과 같은 여러 가지 제약 조건이 있다. 이러한 제약 조건하에서 절감된 조사 예산에 맞춰 대략 추가 심층 조사 가능한 장애인의 규모를 추정해 보았다. <표 8>을 보면 예상되는 전체 장애인 수는 표본조사구 축소 10%인 경우 6,821~7,121명, 20%인 경우 7,260~7,660명, 30%인 경우 7,777~8,277명, 40%인 경우 8,098~8,598명, 50%인 경우 8,416~8,916명으로 나타났다. 표본조사구 축소 비율이 증가할 수록 2017년 실태조사 응답 완료 기준 6,594명보다 더 많은 장애인을 조사할 수 있다. 그러나 조사구를 활용한 조사의 표본 규모가 축소되면 미등록장애인을 구축하는 데 어려움이 발생할 수 있다. 이에 따라 충분하지 않은 미등록장애인의 규모는 대표성 문제를 초래할 수 있고 가중치 산출에도 어려움이 따를 수 있다. 따라서 표본조사구 축소 시 미등록장애인의 특성을 파악하여 많이 분포된 곳은 과대표본추출(oversampling)을 고려해 볼 수 있다.

표 8. 표본조사구 축소 비율에 따른 추가 심층 조사가 가능한 등록장애인 규모

	2017년 조사 기준	표본조사구 축소 비율				
		10%	20%	30%	40%	50%
조사구 수 (개)	2,001	1,798	1,602	1,400	1,203	1,001
가구수 (가구)	36,200	32,520	28,987	25,328	21,767	18,095
A: 전체 장애인 수 (명)	6,820	6,121	5,460	4,777	4,098	3,416
B: 추가 심층 조사가 가능한 등록장애인 수 (명)	-	700~1,000	1,800~2,200	3,000~3,500	4,000~4,500	5,000~5,500
A+B: 예상되는 전체 장애인 수 (명)	-	6,821~7,121	7,260~7,660	7,777~8,277	8,098~8,598	8,416~8,916

자료: 이해정 외. (2020). 전국 단위 실태조사 표본설계 효율화 방안 연구 - 장애인실태조사를 중심으로. 한국보건사회연구원. pp. 8-9. <요약표 1-1>.

5. 나가며

조사데이터는 정책 마련, 정책 효과 평가 등 여러 분야에서 활용되기 때문에 조사데이터의 품질관리는 매우 중요하며, 전 단계(표본설계, 수집, 생산, 배포)에서 지속적으로 해야 한다. 이 글은 표본설계에서의 품질관리, 한국보건사회연구원에서 수행하고 있는 장애인실태조사의 표본설계 효율화 방안과 관련하여 기존 방안(조사구를 활용하는 방안)과 새로운 방안(조사구와 등록장애인 DB를 병행하는 방안)에 대해 장애인 출현율을 산출할 수 있는 범위 내에서 표본조사구를 기존보다 축소할 수 있는지, 표본추출틀을 인구센서스와 더불어 등록장애인 DB도 활용할 수 있는지 등을 중심으로 살펴보았다. 각 방안의 장·단점과 보완 방안은 다음과 같다(표 9).

가. 조사구를 활용하는 방안(현행 방안)

이 방안의 장점은 기존 방식으로 조사를 실시하므로 장애인 출현율을 산출할 수 있을 뿐만 아니라 이전 조사와의 종적 비교도 가능하므로 통계량의 신뢰성도 상대적으로 높은 편이다. 실사관리에서도 그동안의 노하우(knowhow)로 돌발상황 등에 유연하게 대처할 수 있으므로 안정적이라고 볼 수 있다.

다음은 3가지 우려 사항에 대한 내용이다. 첫째, 장애인 여부를 판별하기 위해 1차로 조사하는 가구 및 판별 조사를 대규모로 실시하기 때문에 조사 비용이 많이 소요된다는 점이다. 두 번째, 해가 거듭될수록 조사 환경이 변하여 3만 가구 이상 조사를 완료해야 하는 부담감이 있다는 점이다. 한편, 표본추출틀이 가장 최근의 인구센서스 자료를 활용하지만 시점 차이(2017년 장애인실태조사인 경우 2015년 인구센서스 활용)가

표 9. 조사구를 활용하는 방안 대 조사구와 등록장애인 DB를 병행하는 방안 - 장·단점 및 보완 방안

	조사구를 활용하는 방안	조사구와 등록장애인 DB를 병행하는 방안
장점	<ul style="list-style-type: none"> • 장애인 출현율 산출 가능 • 이전 조사와의 종적 비교가 가능하여 통계량의 신뢰성이 높은 편 • 실사 관리 측면에서 보면 축적된 경험으로 인한 유연한 상황 대처 가능 	<ul style="list-style-type: none"> • 적절한 조사구 규모하에서 장애인 출현율 산출 가능 • 조사구 조사 축소로 인한 가용할 수 있는 조사 예산 확보로, 추가 심층 조사 가능 • 등록장애인 DB의 활용으로 희귀질환 장애유형도 표본에서 누락되지 않고 구축 가능
단점 (우려)	<ul style="list-style-type: none"> • 대규모 가구 및 판별 조사로 많은 조사 비용 소요 • 해가 거듭될수록 조사 환경이 변하여 3만 가구 이상 조사를 완료해야 하는 부담감 • 조사 제반 비용의 인상분이 예산에 반영되지 않는다면 향후 예산 부족 발생 	<ul style="list-style-type: none"> • 2개 추출틀 간 기준 시점 차이로 표본 포함 범위가 다르게 되는 상황 발생 • 표본관리가 어려움 • 기존 조사에 비해 더 많은 연구 인력 필요 및 더 많은 투입 시간 소요
보완 방안	<ul style="list-style-type: none"> • 1상에서 표본 규모 관련 최적값 산출을 위한 심층 연구 필요 • 표본설계 고도화를 위한 등록장애인 DB 활용 방안 마련(표본추출틀 재구축 등) 	<ul style="list-style-type: none"> • 표본조사구 축소 시 미등록장애인 구축 방안 마련(과대표본추출 등)

자료: 이해정 외. (2020). 전국 단위 실태조사 표본설계 효율화 방안 연구 - 장애인실태조사를 중심으로. 한국보건사회연구원. pp. 8-9. <요약표 1-1>.

있다. 이로 인해 재개발, 건물 용도 변경, 신도시 개발 등의 이유로 조사구 변동이 있으므로 조사 진행의 번거로움도 가중될 수 있다. 현재는 관할 동사무소의 협조로 조사가 원활히 수행되고 있는 편이다. 세 번째, 물가 상승에 따른 조사원과 지도원의 인건비, 조사수수료 등과 같은 조사 제반 비용의 인상분이 예산에 반영되지 않는다면 향후 예산 부족 문제로 이어질 수 있다는 점이다.

이 외에 향후 표본설계 고도화를 위하여 등록 장애인 DB 활용 방안을 고려해 볼 수 있다. 그러나 등록장애인 DB의 주소와 조사구의 주소에 대한 매칭(matching)⁵⁾ 작업이 선행되어야 하며, 이는 중요한 과업이 될 것이다. 한편, 인구센서스의 인구 부문 조사 항목에는 활동 제약 문항이 있다. 활동 제약이 있는 사람이면 장애인일 가능성이 클 것이므로 조사구 선정 시 활동 제약 정보를 활용해 볼 수 있다. 현재는 표본설계 시 사용할 수 없는 정보이나 추후 가능하게 된다면 효율적인 표본설계를 구현할 수 있다고 생각한다.

나. 조사구와 등록장애인 DB를 병행하는 방안(새로운 방안)

이 방안의 장점은 인구센서스를 바탕으로 작성된 조사구 조사를 실시하므로 적정한 조사구

규모하에서 장애인 출현을 산출이 가능하다. 또한 등록장애인 DB도 함께 활용하므로 장애 관련 정보 변수(장애유형, 장애등급 등)를 고려한 표본 배분이 가능하여 희귀질환 장애유형도 표본에서 누락되지 않고 구축할 수 있다. 또한 조사구 조사의 규모가 축소되어 가용할 수 있는 조사 예산이 확보되므로 추가로 심층 조사를 할 수 있다.

다음은 우려 사항으로, 첫 번째, 2개의 추출틀 간 기준 시점의 차이로 표본 포함 범위(coverage)가 다르게 되는 상황이 발생할 수 있다는 점이다. 두 번째, 등록장애인 DB는 리스트이므로 동일한 지역으로 묶고 집락을 만들어서 관리하는 등과 관련한 표본관리가 어려울 수 있다는 점이다. 그리고 등록장애인 DB 표본의 경우 등록장애인 DB에서 추출될 확률과 조사구에서 추출될 확률이 중복될 수 있다는 점이 우려된다. 이에 따라 가중치, 신뢰구간 등과 관련한 산출식이 매우 복잡하여 이중추출틀 접근에 의한 추정법의 심층 연구가 필요하다고 생각한다. 세 번째, 이중추출틀을 활용한 조사는 기존 조사에 비해 표본추출에서부터 가중치 산출 관련 통계 업무 및 실사 관리가 기존 조사에 비해 더 많은 연구 인력이 필요하고, 투입 시간도 더 많이 소요된다는 점이다. 그렇기 때문에 충분한 인프라 마련이 선행되어야 할 것이다.

5) '2018년 전국다문화가족실태조사'는 집락의 크기가 일정할 수 있도록 표본추출틀을 재구축하여 조사 비용을 절감하고 표본추출의 효율성을 높이고자 하였다. 즉, 다문화 대상자 리스트를 기반으로 표본설계를 한다면 최종 추출단위를 국적 분류 등으로 직접 총화할 수 있으므로 총화 효율은 높일 수 있다. 그러나 추출된 표본이 산재해서 현장조사가 어렵고 조사 비용이 과다하게 소요될 것이다. 표본추출틀의 재구축 과정은 다음과 같다. 다문화 대상자 명부에 등록된 도로명 주소와 각종 경계 자료(국가기초구역, 인구센서스 통합조사구, 집계구, 도서정보 등)를 공간 연계한다. 이를 바탕으로 최종 조사구 수 기준이 평균 74명 내외가 되도록 행정 구역(시군구) 경계 내에서 인구센서스 조사구를 통합하여 새로운 조사구 경계를 작성하였다(최윤정 외, 2019, p. 727).

현재 실시하고 있는 조사 방법은 여러 가지 상황에 비추어 볼 때 적합한 방법이라고 볼 수 있다. 한편 조사구와 등록장애인 DB를 병행하는 방법의 경우, 다각적으로 검토한 결과 기존 표본 조사구 규모의 30% 이하까지는 축소가 가능하다고 판단한다. 단, 조사의 정확도 및 조사 비용을 종합적으로 고려한 최대허용오차 기준 마련 및 조사구 조사를 통해서만 가능한 미등록장애인의 구축 방안 등에 대해 면밀히 살펴보아야 한다. 등록장애인 DB 정보의 활용과 관련해서는 계속 제고해야 하며, 실제 조사에 바로 적용하기보다는 심층 연구를 통한 신중한 접근이 필요하다고 생각한다. 이 연구 결과는 2017년 장애인실태조사 자료 분석에 한정하였고, 조사 차수별 분석 및 추이를 살펴볼 수 없었던 제약으로 인하여 통계적 분석 결과는 일반화할 수 없다는 점을 밝혀 둔다.

이 글은 장애인실태조사를 중심으로 살펴보았는데, 이러한 특성이 있는 다른 조사에서 표본설계 효율화 방안을 검토할 때 이를 활용해 볼 수 있을 것이다. ■

참고문헌

- 정기원, 권선진, 계훈방. (1995). 1995년도 장애인실태조사. 서울: 보건복지부, 한국보건사회연구원.
- 변용찬, 서동우, 이선우, 김성희, 황주희, 권선진, 계훈방. (2000). 2000년도 장애인실태조사. 서울: 보건복지부, 한국보건사회연구원.
- 변용찬, 김성희, 윤상용, 최미영, 계훈방, 권선진, 이선우. (2006). 2005년도 장애인실태조사. 서울: 보건복지부, 한국보건사회연구원.
- 변용찬, 김성희, 윤상용, 강민희, 손창균, 최미영, 오혜경. (2009). 2008년도 장애인실태조사. 서울: 보건복지부, 한국보건사회연구원.
- 김성희, 변용찬, 손창균, 이연희, 이민경, 이송희,...., 이선우. (2011). 2011년도 장애인실태조사. 서울: 보건복지부, 한국보건사회연구원.
- 김성희, 이연희, 황주희, 오미애, 이민경, 이난희,...., 이선우. (2014). 2014년도 장애인실태조사. 세종: 보건복지부, 한국보건사회연구원.
- 김성희, 이연희, 오욱찬, 황주희, 오미애, 이민경,...., 이선우. (2018). 2017년도 장애인실태조사. 세종: 보건복지부, 한국보건사회연구원.
- 김성희, 이민경, 오미애, 오욱찬, 황주희, 권선진, 오다은. (2019). 2020년 장애인실태조사 사전연구. 세종: 보건복지부, 한국보건사회연구원.
- 손창균, 홍기학, 이기성. (2006). 표본추출 및 관리 매뉴얼. 대전: 통계청, 한국보건사회연구원.
- 이혜정, 오미애, 이민경, 손창균, 박승환, 진재현, 염아림. (2020). 전국 단위 실태조사 표본설계 효율화 방안 연구 - 장애인실태조사를 중심으로. 한국보건사회연구원.
- 최윤정, 김이선, 선보영, 동제연, 정해숙, 양계민,...., 황정미. (2019). 2018년 전국다문화가족실태조사 연구. 서울: 한국여성정책연구원.

〈부표 1〉 총화 기준

연도	정의
2017년	대도시(서울, 부산, 대구, 인천, 광주, 대전, 울산의 동부), 중소도시(경기, 강원, 충북, 충남, 전북, 전남, 경북, 경남, 제주, 세종시의 동부), 농어촌(9개 도 및 세종시 읍·면 지역의 3개 층
2014년	대도시(서울, 부산, 대구, 인천, 광주, 대전, 울산의 동부), 중소도시(경기, 강원, 충북, 충남, 전북, 전남, 경북, 경남, 제주의 동부), 농어촌(9개 도 읍·면 지역의 3개 층
2011년	서울 동부, 광역시 동부, 중소시도의 동부, 전국 읍면부의 4개 층
2008년*	읍면동(집락), 전국 16개 시·도(총화)
2005년	서울의 동, 광역시의 동, 도의 시의 동, 광역시의 읍·면, 도의 시의 읍·면, 도의 군의 읍·면의 6개 층
2000년	서울의 동, 광역시(부산, 대구, 인천, 광주, 대전, 울산)의 동, 기타 시의 동, 읍, 면의 5개 층
1995년	6대 시, 기타 시, 군부의 3개 층

주: 2008년은 1차 추출단위에 대한 정보임.

자료: 이해정 외. (2020). 전국 단위 실태조사 표본설계 효율화 방안 연구 - 장애인실태조사를 중심으로. 한국보건사회연구원. p. 24. 〈표 2-2〉.

〈부표 2〉 시도별 표본조사구 분포 - 2017년 장애인실태조사 기준

지역	동부		읍면부		전체	
	개	%	개	%	개	%
서울특별시	272	19.9	-	-	272	13.6
부산광역시	128	9.3	-	-	128	6.4
대구광역시	80	5.8	-	-	80	4.0
인천광역시	112	8.2	-	-	112	5.6
광주광역시	56	4.1	-	-	56	2.8
대전광역시	80	5.8	-	-	80	4.0
울산광역시	16	1.2	-	-	16	0.8
세종시	8	0.6	8	1.3	16	0.8
경기도	272	19.9	72	11.4	344	17.2
강원도	40	2.9	57	8.9	97	4.8
충청북도	40	2.9	56	8.8	96	4.8
충청남도	32	2.4	72	11.4	104	5.2
전라북도	56	4.1	16	2.5	72	3.6
전라남도	48	3.5	128	20.2	176	8.8
경상북도	48	3.5	120	18.9	168	8.4
경상남도	56	4.1	88	14.0	144	7.2
제주도	24	1.8	16	2.6	40	2.0
전국	1,368	100.0	633	100.0	2,001	100.0

자료: 이해정 외. (2020). 전국 단위 실태조사 표본설계 효율화 방안 연구 - 장애인실태조사를 중심으로. 한국보건사회연구원. p. 56. 〈표 3-2〉.

〈부표 3〉 표본조사구 축소 비율에 따른 지역별 동부·읍면부별 표본조사구 추출 개수

(단위: 개)

지역	표본조사구 10% 축소			표본조사구 20% 축소			표본조사구 30% 축소			표본조사구 40% 축소			표본조사구 50% 축소		
	동부	읍면부	전체	동부	읍면부	전체	동부	읍면부	전체	동부	읍면부	전체	동부	읍면부	전체
서울특별시	245	-	245	218	-	218	190	-	190	163	-	163	136	-	136
부산광역시	115	-	115	102	-	102	90	-	90	77	-	77	64	-	64
대구광역시	72	-	72	64	-	64	56	-	56	48	-	48	40	-	40
인천광역시	101	-	101	90	-	90	78	-	78	67	-	67	56	-	56
광주광역시	50	-	50	45	-	45	39	-	39	34	-	34	28	-	28
대전광역시	72	-	72	64	-	64	56	-	56	48	-	48	40	-	40
울산광역시	14	-	14	13	-	13	11	-	11	10	-	10	8	-	8
세종시	7	7	14	6	6	12	6	6	12	5	5	10	4	4	8
경기도	245	65	310	218	58	276	190	50	240	163	43	206	136	36	172
강원도	36	51	87	32	46	78	28	40	68	24	34	58	20	29	49
충청북도	36	50	86	32	45	77	28	39	67	24	34	58	20	28	48
충청남도	29	65	94	26	58	84	22	50	72	19	43	62	16	36	52
전라북도	50	14	64	45	13	58	39	11	50	34	10	44	28	8	36
전라남도	43	115	158	38	102	140	34	90	124	29	77	106	24	64	88
경상북도	43	108	151	38	96	134	34	84	118	29	72	101	24	60	84
경상남도	50	79	129	45	70	115	39	62	101	34	53	87	28	44	72
제주도	22	14	36	19	13	32	17	11	28	14	10	24	12	8	20
전체	1,230	568	1,798	1,095	507	1,602	957	443	1,400	822	381	1,203	684	317	1,001

자료: 이혜정 외. (2020). 전국 단위 실태조사 표본설계 효율화 방안 연구 - 장애인실태조사를 중심으로. 한국보건사회연구원. p. 63. 〈표 3-6〉, pp. 143-146. 〈부표 A-1〉-〈부표 A-4〉. 재구성

Efficient Sample Design to Improve the Quality of Survey Data - Focusing on the National Survey of Disabled Persons

Lee, Hyejung

(Korea Institute for Health and Social Affairs)

In order to improve the quality of the survey data, we conducted a study for efficient sample design of the National Survey of Disabled Persons.

The National Survey of Disabled Persons covers more than 40,000 households. There are two ways to reduce the burden on a survey based on enumeration districts. The first is to reduce the sample size of household and discrimination surveys. The second is to expand the survey of the disabled based on data other than enumeration districts.

The existing and proposed methods were examined in various ways. The existing method is a survey based on enumeration districts. The proposed method is a survey based on both enumeration districts and databases of individuals registered as disabled.