

International Comparison of Health Care Expenditure

*Existing frameworks,
Innovations and Data Use*

Internationale vergelijking van gezondheidszorguitgaven

*Bestaande kaders, nieuwe
ontwikkelingen en datagebruik*

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
Rector Magnificus

Prof.dr.ir. J.H. van Bommel

en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op
vrijdag 5 september 2003 om 13.30 uur
door

Cornelis Johannes Peter Maria van Mosseveld
Geboren te Dongen

PROMOTIECOMMISSIE

Promotoren Prof.dr. R.M. Lapré
 Prof.dr. F.F.H. Rutten

Overige leden: Prof.dr. J.-P. Poullier
 Prof.dr. W.J. de Gooijer
 Prof.dr. E.K.A. van Doorslaer

8. Een uniforme presentatie van vergelijkbare resultaten kan de uitwisseling van informatie verbeteren tussen analisten en beleidsmakers. Uniformiteit, transparantie en vergelijkbaarheid vereisen permanent aandacht, teneinde (blijvende) bruikbaarheid voor maatschappelijk-economische evaluaties te waarborgen. (W. Brouwer, 1999, *Time and time costs in economic evaluation: taking a societal perspective*)
9. Het toenemend egocentrisme in de maatschappij zal niet alleen leiden tot meer (vraag naar en aanbod van) gezondheidszorg maar tevens tot meer wantrouwen tussen mensen en terughoudend gedrag.
10. Preventieve gezondheidsmaatregelen worden beschouwd als geschikte middelen om de gezondheidsstatus van de bevolking te verbeteren, maar krijgen door het lange termijn perspectief te weinig aandacht van de politiek. (RIVM 2002, *Gezondheid op koers? Volksgezondheid Toekomst Verkenning 2002*)
11. Activiteiten zijn niet alleen een geschikt instrument bij het verbeteren van internationale vergelijkingen, maar ze kunnen ook een belangrijke rol spelen bij het tegengaan van obesitas. (Uitspraak gedaan door Jan Smit tijdens een discussie over hoofdstuk 11 op 4-10-2002)

Publisher

Statistics Netherlands
Prinses Beatrixlaan 428
2273 XZ Voorburg

Printed by

Statistics Netherlands – Facility Services

Cover design

WAT ontwerpers, Utrecht

Information

E-mail: infoservice@cbs.nl

Where to order

E-mail: verkoop@cbs.nl

Internet

<http://www.cbs.nl>

© C.J.P.M. van Mosseveld, Voorburg, 2003.

Voorburg, 2003.

Quotation of source is compulsory. Reproduction is permitted for own or internal use.

The author is fully responsible for the text of this report; the text does not necessarily correspond with the official point of view of Statistics Netherlands.

Key figure: X11

ISBN 903572688 X

Product code: 6008203001

Explanation of symbols

.	= data not available
*	= provisional figure
x	= publication prohibited (confidential figure)
–	= nil or less than half of unit concerned
0 (0,0)	= less than half of unit concerned
–	= (between two figures) inclusive
blank	= not applicable
2002–2003	= 2002 to 2003 inclusive
2002/2003	= average of 2002 up to and including 2003
2002/'03	= crop year, financial year, school year etc. beginning in 2001 and ending in 2002

Due to rounding, some totals may not correspond with the sum of the separate figures.

Acknowledgement

The author would like to express his gratitude for the support of Statistics Netherlands in providing the opportunities to conduct this study. He also owes thanks to his supervising professors Prof.dr. F.F.H. Rutten (of the Erasmus University in Rotterdam) and Prof.dr. R.M. Lapré (Emeritus Professor of the Department of Health Policy and Management of Erasmus University) as well as to Prof.dr. Jean-Pierre Poulhier (Professor at the University of Maastricht and Head of the Department of Health Accounts at WHO Headquarters in Geneva). A special thanks is necessary for Prof. Lapré for offering the author this opportunity and for believing in his abilities and perseverance to complete the research.

In performing the study, the author is very grateful to all his colleagues that helped him, by providing ideas, points of discussion and being sparring partners in the process. He is especially indebted to dr. Gunter Brückner (of Eurostat in Luxembourg), dr. Manfred Huber (of the OECD in Paris) and dr. Markus Schneider (Director of BASYS, Beratungsgesellschaft für angewandte Systemforschung in Augsburg).

The author owes special thanks to Marja Freese for her efforts to get this thesis ready for production at the publishing department.

Jan Smit and Peter van Son deserve a special place in this acknowledgement, because they not only supplied comments and ideas and read through numerous versions of chapters, but also kept him going in difficult times.

It goes without saying that the 'home front' is indispensable in performing any job in a decent way. During the last three years, my wife Brenda and my kids Marinka and Jasper must have suffered from my moods. I thank them deeply for their support and perseverance.

C. J. P. M. van Mosseveld

The Hague

June 2003

Samenvatting

Centrale vraagstelling

De gegevens over de gezondheidszorg, zoals die in diverse nationale en internationale onderzoeken worden gebruikt, zijn niet zonder meer vergelijkbaar over landen heen. Voor veel onderzoekers en gebruikers is de vergelijkbaarheid van dit cijfermateriaal evenwel nauwelijks een punt van aandacht en worden de cijfers gebruikt zoals die worden gepresenteerd.

De hypothese, het uitgangspunt van deze studie, is als volgt geformuleerd:
“De uitgaven aan gezondheidszorg van een geselecteerd aantal landen, zoals weergegeven in diverse bestaande, min of meer geïntegreerde, databestanden, zijn niet vergelijkbaar. Er zijn echter mogelijkheden om de vergelijkbaarheid te vergroten”.

De centrale vraagstelling kan opgedeeld worden in een drietal deelgebieden. Het eerste deelgebied betreft een beschrijving van bestaande systemen en de analyse van de bruikbaarheid van deze systemen voor internationale vergelijkbaarheid van de uitgaven aan gezondheidszorg. Het tweede deelgebied gaat over de beschrijving en analyse van vernieuwingen die ten doel hebben vergelijkbaarheid te vergroten. Het derde deelgebied behandelt het gebruik van data en geeft een analyse van de mogelijkheden van verzoening van bestaande systemen met de vernieuwingen ten behoeve van het vergroten van vergelijkbaarheid.

Bestaande systemen in deze context zijn kaders die sinds de jaren 1950 zijn ontwikkeld en breed worden toegepast in veel landen. Nieuwe methoden in deze studies refereren aan methodologieën en systemen die, hoewel al vijf tot tien jaar bestaand, zich nog steeds in een ontwikkelingsfase bevinden en nog niet op ruime schaal toegepast worden.

Het doel van deze studie is enig licht te werpen op aspecten, die meer en minder belangrijk zijn als het gaat om de vergelijkbaarheid van gegevens over de gezondheidszorg, en daardoor een bijdrage te leveren aan vergelijkbaarheid van uitkomsten van gezondheids(zorg)statistieken en statistieken in het algemeen. Verder wil ik duidelijk maken dat voor de hand liggende conclusies op basis van het bestaande statistische datamateriaal vaak niet in overeenstemming zijn met de realiteit achter de data.

Deel 1: Beschrijving en analyse van bestaande economische kaders

Het eerste deelgebied betreft de hoofdstukken 2 t/m 10. Met betrekking tot vergelijkbaarheid in internationale kaders (hoofdstukken 2 t/m 7) kunnen de volgende vragen worden geformuleerd:

1. Is vergelijkbaarheid wel een relevant onderwerp in bestaande kaders, zoals het Systeem van Nationale Rekeningen (SNR) of het Europees Systeem van Rekeningen (ESR)?
2. Hebben de uitgangspunten en de criteria zoals gehanteerd in het SNR en het ESR consequenties voor de vergelijkbaarheid in het algemeen en de vergelijkbaarheid op het terrein van de gezondheidszorg in het bijzonder?
3. Bieden de ontwikkelingen in nieuwe versies van het SNR en het ESR mogelijkheden tot verbeteringen op het gebied van de vergelijkbaarheid met betrekking tot de economie en de gezondheidszorg?
4. Neemt de mate van vergelijkbaarheid toe indien systemen zoals Satellietrekeningen of Functionele Rekeningen worden gebruikt, die informatie genereren op een meso-economisch niveau, alleen betrekking hebbend op een specifieke branche?

Over de vergelijkbaarheid van de data (hoofdstuk 8 en 9) kunnen de volgende vragen worden gesteld:

5. Beïnvloeden de verschillende methoden waarmee data verzameld en bewerkt worden in het statistische proces de vergelijkbaarheid (van gezondheidszorg-data) tussen landen? Kunnen deze methoden worden veranderd, zodat er een positief effect is op de vergelijkbaarheid?
6. Zijn de kosten van of uitgaven aan gezondheidszorg, zoals weergegeven in diverse databestanden, tussen diverse Europese (EU) landen vergelijkbaar (vergelijkbaarheid over landen)?
7. Is de mate van vergelijkbaarheid van een dataset toegenomen of afgenomen in de tijd in een land (vergelijkbaarheid in de tijd), bijvoorbeeld in Nederland in de periode 1972–1999?

In hoofdstuk 10, handelend over de bestaande kaders en hun mogelijkheden ten aanzien van vergelijkbaarheid van de gezondheidszorguitgaven, worden de conclusies samengevat uit de daaraan voorafgaande hoofdstukken.

Er is helaas geen theorie over vergelijkbaarheid (zie hoofdstuk 2). Wel zijn in diverse artikelen aanzetten gegeven die op termijn zouden kunnen leiden tot een theoretisch kader over vergelijkbaarheid. Tot op heden wordt, om vergelijkbaarheid in beeld te brengen, vaak gebruik gemaakt van gegevens uit bestaande kaders. In de statistiek is een van de meest gebruikte kaders dat van het Systeem van Nationale Rekeningen (SNR). Voor Europese doeleinden is hiervan het Europees Stelsel van Rekeningen (ESR) afgeleid.

Het Systeem van Nationale Rekeningen/Europees Stelsel van Rekeningen (SNR/ESR, zie hoofdstuk 3) heeft als doel de economie, in al zijn relevante aspecten en relaties, macro-economisch volledig en consistent in beeld te brengen. Deze systemen berusten op internationale classificaties en maken gebruik van internatio-

naal overeengekomen definities en afspraken. Macro-economisch (d.w.z. op hoog aggregatieniveau) levert de vergelijking van gegevens tussen landen derhalve betrouwbare resultaten op. Op een lager aggregatieniveau treden evenwel problemen op met de vergelijkbaarheid. Het SNR/ESR is niet gebouwd voor branchevergelijkingen, ook niet voor branches op sociaal gebied, en levert dan ook zelden de informatie op die managers, bestuurders en politici op bijvoorbeeld het terrein van de gezondheidszorg wensen. Desondanks worden bij gebrek aan beter de data uit het SNR/ESR vaak gebruikt. De verbeteringen die in de recente versies van deze systemen zijn aangebracht leveren enige winst op in termen van internationale vergelijkbaarheid, ook van gegevens op brancheniveau. Ten principale echter zijn de gegevens verzameld in het SNR/ESR niet gedetailleerd genoeg. Essentieel hierbij is dat niet te bepalen is of de te vergelijken data berusten op een uniform pakket van activiteiten.

Satellietrekeningen (SR) – zie hoofdstuk 4 – zijn ontwikkeld om een aantal van de nadelen in het gebruik van het SNR/ESR te verhelpen (zoals rigiditeit en overbelasting van het centrale systeem). SR leveren op brancheniveau beter vergelijkbare resultaten op (o.a. door de mogelijkheid meer flexibiliteit in te kunnen brengen en door de grotere mate van detaillering). Het grootste probleem van SR is echter dat de gegevens die worden opgeleverd binnen dit kader nog onvoldoende aansluiten op de gegevens die politici, managers en bestuurders op het terrein van de gezondheidszorg wensen. De relatie met het centrale systeem vormt nog steeds een te stringente voorwaarde. Hoewel de mate van detaillering in een SR kader groter is dan in het SNR/ESR zelf, is ook hier de mate van detaillering van de bronnen van de gegevens niet voldoende om de gewenste vergelijkbaarheid over landen heen te verkrijgen. Ook hier speelt de vraag of de onderliggende activiteitenpakketten, zoals die worden uitgevoerd door de economische entiteiten, uniform zijn.

De Functionele Classificaties (FC) die zijn ontwikkeld binnen het grote kader van het SNR/ESR zijn, net als de SR, een middel om meer flexibiliteit te realiseren binnen het bestaande kader (hoofdstuk 5). FC worden gebruikt om diverse sociale doelen, waaronder gezondheidszorg en onderwijs, beter te laten aansluiten op de eisen van de klanten. FC zijn niet gebonden aan het centrale kader van het SNR/ESR: ook binnen andere kaders kunnen deze classificaties, of onderdelen daarvan, worden gebruikt. De regels die bij implementatie moeten worden gehanteerd, zijn echter voor meer dan één uitleg vatbaar, wat deze classificaties minder bruikbaar maakt voor internationale vergelijking. Echter, zelfs indien besloten zou worden eenheden te laten fungeren als het centrale uitgangspunt in deze classificaties, houdt dit niet in dat de combinatie van activiteiten die ten grondslag ligt aan deze eenheden tussen landen (nagenoeg) identiek is.

Gezondheids(zorg)rekeningen (hoofdstuk 6) worden in veel landen gemaakt om de informatie te kunnen genereren die door beleid en bestuur wordt verlangd. Deze

rekeningen zijn over het algemeen nog gedetailleerder dan de reeds eerder genoemde opstellingen zoals Nationale Rekeningen (NR) en Satellietrekeningen (SR). Gezondheids(zorg)rekeningen zijn in principe niet vergelijkbaar over landen heen, omdat dit soort opstellingen voor een nationaal doel wordt opgezet. Gezondheids(zorg)rekeningen geven derhalve geen vergelijkbare uitkomsten, maar geven inzicht in de verschillen tussen landen en de gebruikte systemen van zorg. Uitsluitend specifieke vraagstellingen, bedoeld om internationale vergelijkingen uit te voeren, kunnen enige verbetering in de vergelijkbaarheid van de uitkomsten van gezondheids(zorg)rekeningen opleveren. Het basisprobleem bij alle internationale vergelijkingen, de samenstelling van het activiteitenpakket van de eenheden of programma's die worden vergeleken, wordt ook met het samenstellen van nationale gezondheids(zorg)rekeningen (zelfs indien gebruik gemaakt wordt van een standaardset definities en classificaties) niet opgelost.

Wat betreft de uitgaven aan de gezondheidszorg is één van de beste databases de database die onderhouden wordt door de OESO (Organisatie voor Economische Samenwerking en Ontwikkeling), de OECD Health Data Base (hoofdstuk 7). Alle landen die lid zijn van de OESO leveren hiervoor gegevens. Hoewel door het secretariaat van de OESO grote moeite wordt gedaan alle gegevens zo vergelijkbaar mogelijk te krijgen, zeker de gegevens betreffende de uitgaven aan de gezondheidszorg, is gebleken dat volledige vergelijkbaarheid over landen heen na ruim 25 jaar nog steeds niet bereikt is. Grote problemen ontstaan door verschillen in terreindekking, definities van termen, kwaliteit van de door landen verzamelde data en gehanteerde berekeningsmethoden. Het belangrijkste probleem dat schuil gaat achter de gegevens in de database is echter, dat de inhoud van de eenheden die vergeleken worden ten principale niet identiek is. Eenheden met dezelfde begripsaanduiding in diverse landen hebben niet per definitie dezelfde inhoud, zeker niet indien gekeken wordt naar de door deze eenheden uitgevoerde activiteiten.

Van de basis statistische processen (zie hoofdstuk 8) die kunnen worden onderscheiden (input, throughput, output), kan alleen het throughput proces door de statistiekmakers volledig worden gecontroleerd. De productie van cijfers voor internationaal gebruik heeft evenwel in veel statistiekbureaus geen hoge prioriteit, tenzij het data betreft die politiek van nationaal en internationaal belang zijn (bijv. data vereist voor de Europese Monetaire Unie, EMU). Data op het gebied van de gezondheidszorg behoren hier (nog) niet toe.

Vergelijkbaarheid van de data (hoofdstuk 9) staat niet a-priori vast. Echter indien data gebaseerd op een solide systeem als het Systeem van Nationale Rekeningen/Europees Stelsel van Rekeningen (SNR/ESR) worden vergeleken, kan worden geconcludeerd dat, hetzij de data niet vergelijkbaar zijn, hetzij de gebruikte methoden niet identiek zijn. Welke bron ook gebruikt wordt om vergelijkingen van de zorg tussen landen op te baseren (SNR/ESR of OESO), op voorhand is niet te

bepalen welke bron de beste vergelijking oplevert. Nationale rekeningen data zijn niet per definitie beter of slechter voor internationale vergelijkingen dan data verzameld door de OESO in de Health Data Base.

Vergelijkingen van de uitgaven aan de gezondheidszorg in de tijd binnen een land leveren alleen bruikbare uitspraken op, indien naar belangrijke beleidsmaatregelen (zowel gezondheidszorgspecifiek als meer algemeen economisch) wordt gekeken. Hierbij moet opgemerkt worden, dat het welhaast onmogelijk is de effecten van één afzonderlijke maatregel te isoleren. Indien beleidsmaatregelen alleen de gezondheidszorg betreffen, is bovendien een grote mate van kennis van het terrein (statistisch en beleidsmatig) nodig en een grote hoeveelheid detailinformatie. Bij soortgelijke vergelijkingen tussen landen is uiteraard dezelfde hoeveelheid kennis en informatie voor deze landen nodig, willen er bruikbare conclusies kunnen worden getrokken.

Uitspraken over vergelijkbaarheid kunnen eigenlijk pas valide gedaan worden indien internationaal exact bekend is, waar de grenzen van het onderzoeksterrein zijn gesitueerd en welke actoren/eenheden zijn opgenomen binnen de grenzen van de (gezondheids)zorg en welke (clusters van) activiteiten door deze actoren/eenheden worden uitgevoerd.

Uit de analyse van de bestaande kaders en hun mogelijkheden voor internationale vergelijkbaarheid (zie hoofdstuk 10) is een tweetal opmerkelijke constatering af te leiden:

- Op de eerste plaats kan geconstateerd worden, dat vergelijkbaarheid een relatief begrip is: bepalend zijn het object dat vergeleken moet worden en het doel dat met de vergelijking wordt nagestreefd. Elk doel vereist een eigen vergelijkingskader.
- Een tweede constatering is dat bestaande kaders niet de vergelijkingsmogelijkheden bieden die over het algemeen door de gebruikers van gezondheidszorggegevens worden gevraagd. Bestaande kaders, zoals Nationale Rekeningen, bieden naast consistentie en samenhang in de verwerking van het datamateriaal, tevens mogelijkheden tot vergelijkbaarheid op een hoog aggregatieniveau. Op een lager niveau van aggregatie, bijvoorbeeld het niveau van bedrijfstakken, wijkt het doel van de bestaande kaders af van het doel internationale vergelijkbaarheid. Bovendien hebben de objecten die vergeleken worden een onvergelijkbare inhoud. Activiteiten zijn nodig om op dit lagere detailniveau vergelijkbaarheid mogelijk te maken.

Deel 2: Vernieuwingen ten behoeve van internationale vergelijkbaarheid

In het tweede deel van deze studie (hoofdstukken 11 t/m 14) wordt een aantal vernieuwingen gepresenteerd, die de vergelijkbaarheid van de data over de gezondheidszorg kunnen vergroten. Over de mogelijkheden om de vergelijkbaarheid te vergroten, kunnen de volgende vragen worden gesteld:

8. Er zijn recent, in internationaal verband, instrumenten ontwikkeld die de mate van vergelijkbaarheid kunnen vergroten, zoals het Gemeenschappelijk Vergelijkbaar Pakket en het Systeem van Gezondheidsrekeningen. Is er overeenstemming over het feit dat deze instrumenten aanknopingspunten bieden voor verbetering van de internationale vergelijkbaarheid?
9. Welk perspectief is er om deze nieuwe instrumenten samen te brengen in een internationaal acceptabel kader (zoals het Systeem van Nationale Rekeningen of het Europees Stelsel van Rekeningen)?

Mogelijkheden worden gepresenteerd om de objecten van vergelijking een meer identieke inhoud te geven, zodat de vergelijking tussen landen in ieder geval niet gehinderd wordt door inconsistenties in het basismateriaal (de objecten van vergelijking).

Een universeel kenmerk van alle bestaande kaders is de activiteit, die in economische zin wordt uitgeoefend door eenheden. De wijze waarop met deze activiteiten wordt omgegaan, is in de diverse benaderingswijzen en kaders verschillend. Ook in de verbetertrajecten die in dit onderzoek worden besproken, wordt op een eigen wijze met activiteiten omgegaan.

Het eerste verbetertraject dat besproken wordt (in hoofdstuk 11), is de methode van het Gemeenschappelijk Vergelijkbaar Pakket (GVP). In deze methode wordt sterk ingezet op de éénduidigheid van activiteiten als noodzakelijke voorwaarde om de vergelijkbaarheid tussen landen te vergroten. Aan de basis van deze methode staat, dat alle eenheden in het economisch verkeer gekenmerkt worden door de activiteiten, die eenheden uitvoeren. Activiteiten zijn echter niet de dragers van informatie. Informatie over de activiteiten wordt verstrekt via eenheden, o.a. verstrekkers van zorg. In dit GVP-onderzoek, handelend over de gezondheidszorg, is, bij gebrek aan een voldoende gedetailleerde en geharmoniseerde activiteitenclassificatie, een lijst met activiteiten ¹⁾ ontwikkeld. Aangezien het vertrekpunt van deze methode ligt in de activiteiten zelf, is een enorme hoeveelheid informatie nodig, om alle eenheden voldoende gekarakteriseerd te krijgen. Met behulp van een netwerk van experts is deze informatie verzameld. In de methode van het GVP worden daarom, in eerste instantie via een bilaterale vergelijking tussen twee landen, eenheden op een zodanige wijze geclusterd, dat de inhoud van het totale pakket van activiteiten van beide landen (nagenoeg) identiek wordt. In een vergelijking, waarbij N landen zijn betrokken, levert deze procedure N-1 Bilaterale Vergelijkbare Pakketten (BVP) op, indien één van de landen als sleutelland wordt aangemerkt. Voor een Gemeenschappelijk Vergelijkbaar Pakket, geldig voor alle landen in de vergelijking, moeten de verschillende Bilaterale Pakketten geüniformeerd worden tot één GVP. Dit GVP bleek in de praktijk geconstrueerd te kunnen worden. Bovendien is vastgesteld dat dit GVP stabiel van omvang was, in die zin dat de toevoeging van landen slechts marginale invloed op de inhoud, de omvang van het pakket van activiteiten, zou hebben. Om deze methode toe te passen is echter een grote hoeveelheid zeer

gedetailleerde informatie nodig, alsmede een grote kennis van de te vergelijken systemen van gezondheid(szorg).

Het voordeel van het GVP is, dat er internationaal consensus is over de methode die aan het GVP ten grondslag ligt: het gebruik van activiteiten als criterium voor selectie op vergelijkbaarheid. De geconstateerde nadelen van deze methode hebben betrekking op enerzijds de grote mate van bewerkelijkheid om vergelijkbaarheid in termen van uniforme activiteitenpakketten mogelijk te maken, de hoeveelheid benodigde informatie, en het ontbreken van (een link naar) een theoretisch kader.

Het Systeem van Gezondheidsrekeningen (SGR) wordt besproken in hoofdstuk 12. Dit systeem, ontwikkeld door de OESO en gesteund door de EU, gaat in principe van dezelfde uitgangspunten uit. Aangezien echter de mate van detail en kennis nodig om het GVP tot een succes te maken, alsook de hoeveelheid werk die nodig is, voor een internationale organisatie niet haatbaar zijn op jaarbasis, is een andere invalshoek gekozen. Ook in deze benadering vormen de uitgevoerde activiteiten en de eenheden het uitgangspunt. Eenheden op het gebied van de gezondheidszorg in de verschillende landen zijn niet vergelijkbaar, activiteiten wel. In deze methode worden echter niet de eenheden geclusterd tot vergelijkbaarheid aanwezig is, maar worden activiteiten geclusterd tot hanteerbare uniforme clusters, functies genoemd. Aangezien deze functies door alle eenheden in variabele samenstellingen worden uitgevoerd, leveren uiteindelijk de functies de vergelijkingsgrond binnen het SGR. In het SGR zijn diverse classificaties ontwikkeld (classificatie van functies, classificatie van financiers en een classificatie van verstrekkers van gezondheidszorg). Het gebruik van deze methode levert idealiter een drie-assig systeem op, waarbij de gegevens in hun onderlinge verband kunnen worden gepresenteerd. Binnen het SGR is speciale aandacht geschonken aan de mogelijkheden, om deze classificaties te koppelen aan de classificaties van de Nationale Rekeningen. De voordelen van het SGR hebben betrekking op de aanwezigheid van een theoretische onderbouwing en de sterke inbedding in het Systeem van Nationale Rekeningen. De nadelen van het SGR betreffen vooral de compromissen die tot stand gekomen zijn bij de samenstelling van de classificatie van functies. Aan de zuiverheid van functies is onvoldoende recht gedaan, terwijl bovendien een providerkenmerk (de productiemodus) onterecht gemixt is met functiekenmerken.

Het Nederlandse Systeem van Zorgrekeningen (zie hoofdstuk 13) – de opvolger van de statistiek ‘Kosten en Financiering van de Gezondheidszorg in Nederland’ – is gebaseerd op het Systeem van Gezondheidsrekeningen (SGR). In de toepassing van de systematiek van het SGR zijn verschillende gebruiksdoelen geformuleerd, waarvoor verschillende classificaties van eenheden het uitgangspunt zijn. Gedurende het proces van opzetten en implementeren van dit nieuwe statistische systeem zijn de eerder genoemde onvolkomenheden in het SGR aan het licht gekomen. Verbeteringen ten opzichte van het SGR zijn doorgevoerd met betrekking tot de zuiverheid

van de gehanteerde functies. De onjuiste toevoeging in het SGR van productiemodi aan de functies in plaats van aan de verstrekkers van de zorg is ongedaan gemaakt. In het Nederlandse Systeem van Zorgrekeningen worden gegevens gepresenteerd, zowel uitgaande van de OESO-classificaties, als van de Nederlandse versies van deze classificaties.

Tenslotte worden zowel de Classificatie van Functies (ICHA-HC) als de Classificatie van Gezondheidszorgverstrekkers (ICHA-HP) uitgebreid met functies op sociaal gebied en actoren in de sociale sector. De Classificatie van Gezondheidszorgverstrekkers wordt daarvoor omgebouwd tot een Classificatie van Gezondheids- en Sociale Zorgverstrekkers (HSP).

In een volgend hoofdstuk (hoofdstuk 14) worden kort andere mogelijkheden, om gegevens op het gebied van de gezondheidszorguitgaven meer vergelijkbaar te kunnen krijgen, besproken.

Een ingang kan bijvoorbeeld zijn de Internationale Classificatie van Ziekten (ICD) ontwikkeld door de Wereld Gezondheidsorganisatie WHO. Hoewel interpretatieverschillen in het gebruik van deze classificatie mogelijk zijn, kunnen uitgaven naar ICD of ICD-cluster een verbetering in de vergelijkbaarheid bieden. Aangezien uitgaven zijn gekoppeld aan verstrekkers, is de problematiek met betrekking tot de identificatie van eenheden gelijk aan de hiervoor besproken methoden. Gesamenlijk over de totaliteit van ICD-codes levert deze methode goed vergelijkbare gegevens, indien de inhoud van de eenheden waarop de uitgaven betrekking hebben over landen heen identiek is.

Voor alle andere mogelijkheden die gebruikt kunnen worden (zoals ESSPROS, Europees Systeem van geïntegreerde Statistieken op het gebied van Sociale Bescherming, DRG, Diagnose Gerelateerde Groepen), is de problematiek niet anders. In alle gevallen moeten variabelen (zoals uitgaven) worden gelinkt aan andere variabelen (ziekten, functies, etc.). Vergelijkbaarheid van de uitkomsten is afhankelijk van de vergelijkbaarheid van de eenheden, die het uitgangspunt vormen.

De weg die is ingeslagen, met het in Nederland door het CBS (Centraal Bureau voor de Statistiek) ontwikkelde Gemeenschappelijk Vergelijkbaar Pakket en voortgezet door de OESO (Organisatie voor Economische Samenwerking en Ontwikkeling) met het Systeem van Gezondheidsrekeningen (SGR), levert een duidelijke verbetering op van de vergelijkbaarheid. Het probleem van vergelijkbaarheid, zelfs van gezondheidszorguitgaven, is echter nog niet geheel opgelost. Er is nog veel theoretisch en praktisch werk nodig om verdere verbeteringen in de vergelijkbaarheid te verkrijgen.

De inspanningen, die getoost moeten worden om bijvoorbeeld een SGR op te zetten moeten niet onderschat worden. Het doel is dat de inspanningen die gedaan worden op termijn hun vruchten zullen afwerpen, via een betere internationale vergelijkbaarheid. Hoewel binnen de Europese Unie het sociale beleid het prerogatief van

de individuele regeringen blijft, zullen de afzonderlijke lidstaten alsook de EU als geheel gebaat zijn bij een betere vergelijkbaarheid van gegevens over de gezondheidszorg.

Deel 3: Data: gebruik en verzoening

Het derde deelgebied (hoofdstukken 15 en 16) betreft het gebruik van data en de verzoening van de data uit de besproken kaders en de vernieuwingen ten behoeve van de vergelijkbaarheid. Data worden pas relevant indien die gebruikt worden. De volgende deelvragen kunnen hierover worden geformuleerd.

10. Welke data worden gebruikt in een nationale en een internationale context? Wordt het beseft dat data niet altijd vergelijkbaar zijn, welk besef de conclusies beïnvloedt die gebaseerd zijn op de uitkomsten? Wordt het beseft dat een zekere kwantiteit en kwaliteit van data nodig zijn voor onderzoek?
11. Tenslotte is het de vraag of de data uit de verschillende systemen die gebruikt worden in internationale vergelijkingen van gezondheidsuitgaven met elkaar te verzoenen zijn. En wat moet er veranderd of toegevoegd worden om een acceptabel niveau van vergelijkbaarheid te verkrijgen?

In hoofdstuk 15 wordt het gebruik van de data centraal gesteld, zowel nationaal als internationaal. Nationaal wordt door beleidsmakers veelvuldig gebruik gemaakt van data die in internationale databases aanwezig zijn. In nogal wat gevallen wordt vergeten de mate van vergelijkbaarheid van de data te vermelden. Ook in internationale artikelen en econometrisch onderzoek over de gezondheidszorg wordt hiervan niet altijd melding gemaakt. Vooral de mate waarin conclusies worden beïnvloed wordt veelal niet beseft. Een zeer belangrijk item bij gebruik van data voor onderzoek is de hoeveelheid en de kwaliteit van de data. Een zekere (grote) mate van kwantiteit en kwaliteit is een voorwaarde om valide onderzoek te kunnen uitvoeren.

Vervolgens worden in hoofdstuk 16 de mogelijkheden onderzocht om de data met betrekking tot gezondheidszorguitgaven in de bestaande systemen met elkaar in overeenstemming te brengen. Deze test wordt uitgevoerd met Nederlandse data, omdat die in voldoende gedetailleerde mate aanwezig zijn. Vervolgens worden de gevolgen bekeken van de resultaten van deze nationale synthese om de behoefte aan internationaal vergelijkbare gegevens over de uitgaven aan gezondheidszorg te kunnen bevredigen. Vergelijkbaarheid wordt afgemeten aan het internationaal opgezette Systeem van Gezondheidsrekeningen, aangepast aan de Nederlandse eisen met betrekking tot de classificatie van functies. Een problematisch punt is evenwel de benodigde mate van detail in data gebaseerd op de gehanteerde classificaties (vooral de classificatie van activiteiten maar ook de classificaties van goederen en diensten en die van functies).

Op basis van NR-data (Nationale Rekeningen) blijkt dat, afhankelijk van het uitgangspunt (productie of consumptie) en de gehanteerde criteria, niet meer dan een beperkt niveau van vergelijkbaarheid kan worden bereikt. Naar mijn mening is de mate van vergelijkbaarheid die gerealiseerd kan worden met behulp van andere methoden dan gezondheidsrekeningen niet voldoende.

In hoofdstuk 17 worden de conclusies over de vernieuwingen op het gebied van de internationale vergelijkbaarheid en het gebruik en de verzoening van data gepresenteerd.

Uit de analyse van de vernieuwende ontwikkelingen, hun mogelijkheden voor internationale vergelijkbaarheid en het gebruik van data zijn de volgende opmerkelijke constatering af te leiden:

- Vergelijkbaarheid kan alleen bereikt worden door activiteiten te gebruiken. Activiteiten zijn een leidend beginsel in elke poging om de vergelijkbaarheid van gezondheidszorguitgaven tussen landen te vergroten.
- Vergelijkbaarheid tussen landen kan verder vergroot worden door een classificatie specifiek voor dit doel te ontwikkelen. De Classificatie van Functies (in een zuivere vorm) kan zo'n instrument zijn.
- Bij het gebruik van data in internationaal vergelijkend onderzoek zijn meer opmerkingen over de mate van (on)vergelijkbaarheid van de data nodig. Vooral de invloed van het gebrek aan vergelijkbaarheid op de uitkomsten moet beter gespecificeerd worden. Het gebruik van data uit bestaande systemen, vooral NR, geeft slechts een beperkte mate van bruikbaarheid. Vooral omdat de aansluiting met de nieuw ontworpen Classificatie van Functies slechts met grote inspanningen kan worden gemaakt.

Aanbevelingen

Voor de toekomst kunnen de volgende verbeteringen met betrekking tot vergelijkbaarheid van gezondheidszorgdata (zie hoofdstuk 18) worden bereikt. Hierbij kan worden gedacht aan:

- *Activiteiten*. Aangezien activiteiten in alle kaders (bestaande zowel als nieuw ontwikkelde) een cruciale rol spelen, is het belangrijk, dat er een volledige internationaal geaccordeerde classificatie van activiteiten komt, ingebed in de classificatie van functies. Een eerste aanzet voor zo'n classificatie van activiteiten op het gebied van de gezondheidszorg is gegeven in het Gemeenschappelijk Vergelijkbaar Pakket (GVP)-project. Vervolmaking van deze lijst van activiteiten en uitbreiding naar activiteiten in de sociale zorg, alsmede de koppeling van deze activiteiten met bestaande classificaties, is nodig.
- *Functies*. Omdat uniforme functies (als clusters van activiteiten) in de opzet van het Systeem van Gezondheidsrekeningen (SGR) van groot belang zijn, moet grote

nadruk worden gelegd op de zuiverheid van deze functies, evenals op de wederzijdse uitsluitbaarheid ervan. Vermenging van functieonderdelen onderling en vermenging met andere aspecten, zoals de productiemodus, moet via internationaal overleg uitgesloten worden.

- *Integratie van branches.* Omdat de scheidslijnen tussen de terreinen van gezondheidszorg en sociale zorg steeds meer vervagen, wordt de integratie van deze twee terreinen voorgesteld. Het opnemen van clusters van verstrekkers van sociale zorg in de classificatie van verstrekkers moet overwogen worden. Uitbreiding van de classificatie met de sociale sector moet nader onderzocht en vervolmaakt worden.
- *Verbindingen met Nationale Rekeningen.* Het spreekt vanzelf dat verbindingen met het internationaal geaccordeerde systeem van Nationale Rekeningen gelegd moeten worden, willen nieuwe systemen op deelreinen bestaansrecht hebben. Op welke wijze deze verbindingen gelegd moeten en kunnen worden (topdown vanuit de geaggregeerde systemen van het Systeem van Nationale Rekeningen (SNR) of bottom-up vanuit de detailinformatie van bijvoorbeeld het SGR) is een onderwerp van verder onderzoek.
- *Productenclassificaties.* De CPA/CPC (Statistische Classificatie van Producten naar Activiteit in de Europese Unie/Centrale Product Classificatie) kunnen bruikbaar zijn in de vergelijking van de gezondheids- en welzijnszorguitgaven tussen landen, als de classificaties volledig in al hun details zouden worden gebruikt. Echter, om volledige vergelijkbaarheid te krijgen moeten de classificaties worden verdiept en meer gedetailleerd worden op het gebied van de gezondheids- en welzijnszorg. Aangrijpingspunten voor deze verdieping kunnen gevonden worden in de lijst van activiteiten zoals die is samengesteld voor het GVP-project.

Volledige vergelijkbaarheid is hiermee nog niet bereikt. Samenwerking en coördinatie tussen alle betrokken organisaties en hun leden is nodig, teneinde verdere verbeteringen te realiseren.

Noot:

- 1) In de bijlagen bij hoofdstuk 11 over het GVP zijn de eerste resultaten weergegeven van koppelingen tussen de lijst met activiteiten en de classificatie van functies van de OECD (ICHA-HC) alsmede de classificatie van goederen (CPA).

Afkortingen (gebruikt in de Samenvatting)

BVP	Bilateraal Vergelijkbaar Pakket
CBS	Centraal Bureau voor de Statistiek
GVP	Gemeenschappelijk Vergelijkbaar Pakket
CPA	Statistische Classificatie van Producten naar Activiteit in the Europese Economische Gemeenschap
CPC	Centrale Producten Classificatie
DRG	Diagnose Gerelateerde Groep
EMU	Europese Monetaire Unie
ESR	Europees Systeem van Rekeningen
ESSPROS	Europees Systeem van geïntegreerde Statistieken over Sociale Bescherming
EU	Europese Unie
FC	Functionele Classificaties
HSP	Gezondheidszorg- en Sociale Zorgverstrekkers
ICD	Internationale Classificatie van Ziekten
ICHA	Internationale Classificatie voor Gezondheidsrekeningen
ICHA-HC	Internationale Classificatie voor Gezondheidsrekeningen – Classificatie van Functies in de Gezondheidszorg
ICHA-HP	Internationale Classificatie voor Gezondheidsrekeningen – Classificatie van Verstrekkers van Gezondheidszorg
NR	Nationale Rekeningen
OESO	Organisatie voor Economische Samenwerking en Ontwikkeling
SC	Sociale Zorgfuncties
SGR	Systeem van Gezondheidsrekeningen
SNR	Systeem van Nationale Rekeningen
SP	Sociale Zorgverstrekkers
SR	Satellietrekeningen
WHO	Wereld Gezondheidsorganisatie

Summary

Central question

Data on health care, as used in various national and international research projects, cannot be compared between countries just like that. Many researchers and users pay hardly any attention to the comparability of these data and use them just as they are presented.

My hypothesis, and the starting point of this study, is formulated as follows:

“Expenditure on health care of a selection of countries as presented in various, more or less integrated existing databases, is not comparable. However, there are possibilities to enlarge the comparability of these data”.

The central question can be divided in three areas. The first concerns a description of existing economic frameworks and the analysis of the usefulness of these systems for cross-country comparability of health care expenditure. The second deals with the description and analysis of innovations aiming to enlarge comparability. The third area covers the use of data and the reconciliation of the existing frameworks with the innovations regarding possibilities to enlarge comparability.

Existing systems in this context are frameworks developed since the 1950's and broadly applied in many countries. New methods in this study refer to methodologies and systems, which, although they have existed for five to ten years, are still in a developing stage and not yet broadly applied.

The goal is to shed some light on aspects that are to a greater or lesser extent, important when comparing data on health care, and thus to provide a contribution to comparability of the results of health (care) statistics and statistics in general. Furthermore, I will make clear that obvious conclusions based on statistical material are often not in accordance with reality behind the data.

Part 1: Description and analysis of existing economic frameworks

Chapters 2 to 10 cover the first area. With respect to comparability in international frameworks (chapters 2 to 7), the following questions can be posed:

1. Is comparability an issue in existing international frameworks like the System of National Accounts (SNA) and the European System of Accounts (ESA)?
2. Do the starting points and criteria mentioned in for example the SNA and ESA hamper or promote comparability, especially comparability of health care data?

3. Does the development of new versions of the SNA and ESA lead to improvements in comparability in general terms, and in health care terms in particular?
4. Does the level of comparability increase when more detailed systems like satellite accounts or (functional or national) accounts are used, which generate information on a meso-economic level dealing only with a specific branch?

On the topic of comparability of the data (chapters 8 and 9) the following questions are posed:

5. Do the various ways data are collected and treated within the statistical process influence cross-country comparability (of health care) and if so can this influence be changed so that it is positive for comparability of health care data?
6. Are the data on health care cost/expenditure of a selection of European (EU) countries as presented in various databases comparable (comparability across countries)?
7. Has the amount of (in)comparability of a set of data elements increased or decreased over time (time comparability) within a country: for example the Netherlands in the period 1972–1999?

Chapter 10 summarises conclusions from the previous chapters on existing frameworks and their possibilities for comparability of health care expenditure

Unfortunately, there is no theory on comparability (see chapter 2). Various articles contain initiatives that might in due course result in a theoretical framework of comparability. Until now, only the existing frameworks can be used to map cross-country comparability. Statistically one of the most frequently used frameworks is the System of National Accounts (SNA), and for European purposes, the derived European System of Accounts (ESA).

The goal of SNA/ESA (chapter 3) is to present the economy in all its relevant aspects and relations, complete and consistent in a macro-economic sense. The systems are based on international classifications and use internationally agreed definitions and conventions. Macro-economically (i.e. at a high level of aggregation), comparison of data across countries yields reliable results. Comparability problems start occurring at a lower level of aggregation. The SNA/ESA is not built for comparisons by branch, certainly not for branches in the social area. The SNA/ESA rarely supplies the information desired by managers, administrators and politicians in the area of health care. Nevertheless, for want of better information, data originating from SNA/ESA are frequently used. Improvements introduced in the recent versions of these systems have led to some gains in terms of cross-country comparability, also at the branch level. In principle, however, the data collected in the SNA/ESA are not detailed enough. Essentially it is not known whether the data to be compared are based on a uniform package of activities.

Satellite accounts (SA) – see chapter 4 – have been developed to remove a number of disadvantages in using the SNA/ESA (e.g. rigidity and overburdening of the central framework). SA provide better comparable results at the branch level (among other things through the possibility of introducing more flexibility and the larger amount of detail). The largest problem of SA, however, is still that the data supplied within this framework are not the data needed by politicians, managers and administrators in the area of health care. Probably the relation with the central framework is too stringent a condition to be met. Although the level of detail within an SA framework is larger than in the SNA/ESA, the level of detail of the sources of the data is not sufficient to result in the desired cross-country comparability. Again the question of the uniformity of the underlying packages of activities, as performed by the economic entities, is relevant.

Like SA, the functional classifications (FC), developed within the large framework of SNA/ESA are intended to create more flexibility within existing frameworks (Chapter 5). FC are used to improve the connections of various social goals, among which health care, with the demands of customers. Of course, FC are not tied to the central framework of SNA/ESA; they or parts of them can be used in other frameworks, too. However, the rules to be used in implementing them can be interpreted in more than one way, which makes these classifications less suitable for cross-country comparisons. Even if (economic) units are decided on as the core starting point in this classification, this does not imply that the combination of activities underlying these units is (almost) identical across countries.

Many countries construct health (care) accounts (chapter 6) to be able to generate the information requested by policy and management. These accounts are usually more detailed than the aforementioned national and satellite accounts. Health (care) accounts are in principle not comparable across countries, because they are created for a national purpose. Consequently, health (care) accounts do not supply comparable results, but provide insight into the differences between countries and the systems of care used. Only specific questionnaires aimed at cross-country comparisons can provide some improvements in the comparability of the results of these accounts. The basic problem in all international comparisons, i.e. the composition of the activities of the units and or programmes compared, is not solved in the construction of national health (care) accounts, not even if a standard set of definitions and classifications is used.

For health care-related expenditure, one of the best databases is that maintained by the OECD (Organisation for Economic Co-Operation and Development), the OECD health database (chapter 7). All member countries of the OECD supply information for this database. Although the Secretariat puts considerable effort in making the data as comparable as possible, especially data on health care, it is apparent that complete comparability has not been achieved after a period of 25 years. There are

substantial problems concerning differences in the area covered, definitions of terms, quality of the data collected by countries and calculation methods used. The main problem hidden behind the data in the database is, however, that the contents of the units compared are in principle not identical. Units with the same name in various countries do not have the same contents, certainly not on an activity basis.

Of the three basic distinguished statistical processes (see chapter 8) – input, throughput, and output – only the throughput process can be controlled completely by statisticians. The production of data for international use does not have a high priority in many statistical offices, unless they are of national or international political importance (e.g. data required for the European Monetary Union, EMU). Data in the area of health care are not (yet) in this sphere.

Comparability of the data (chapter 9) is not determined beforehand. However, if data based on a solid system like SNA/ESA are compared, it can be concluded that either the data are not comparable or the methods used are not identical. Whichever source is used to compare care across countries, it is unknown beforehand which source provides the best comparison. National accounts are not by definition better or worse for cross-country comparisons than data collected by the OECD in the health database.

Comparisons of expenditure of health care in time for a single country provide only usable judgements if important policy measures (both health care-specific and the more general economic measures) are taken into consideration. It should be noted that it is hardly possible to isolate the effect of a single measure. Moreover, if measures relate only to health care, a wide statistical and policy knowledge of the area is necessary, as well as a large amount of detailed information. It goes without saying that in comparisons of this kind across countries the same amount of knowledge and information is needed for these countries to be able to draw useful conclusions.

Valid conclusions on comparability can only be based on exact knowledge of the international boundaries of the area under study, and knowledge of which actors/units are included within the boundaries of (health) care and which (clusters of) activities these actors/units perform.

Two remarkable conclusions can be drawn from the analysis of the existing frameworks and their possibilities for cross-country comparisons (see chapter 10) are:

- First it can be concluded that comparability is a relative concept: determining factors are the object to be compared and the goal pursued with the comparison. Every goal demands its own framework of comparison.
- A second conclusion is that existing frameworks do not offer possibilities for comparison that are demanded in general by the users of health care data. Existing frameworks like national accounts offer comparability at a high level of

aggregation, alongside consistency and coherence in the processing of the data. At a lower level of aggregation (e.g. at branch level), the goal of existing frameworks deviates from the goal of cross-country comparability. Furthermore, objects contain non-comparable contents. Activities are necessary at this lower level of aggregation to make comparisons possible.

Part 2: Innovations concerning cross-country comparability

The second part of this study (chapters 11 to 14) presents a number of innovations that can enlarge the comparability of health care data. With respect to the possibilities for improvement of the comparability, the following questions can be formulated:

8. Recently a number of instruments have been developed that increase the level of comparability, like the *Common Comparable Package* approach and the *System of Health Accounts*. Is there any agreement on the possibilities these instruments offer in improving cross-country comparability?
9. Which opportunities are available to enclose all these new methods in an internationally acceptable framework (like the SNA or ESA)?

Possibilities are presented, through which the objects of comparison receive more identical contents, so that at least the goals that are pursued with cross-country comparisons are not hampered by inconsistencies in the basic material (the objects of comparison).

A universal feature of all existing frameworks is the activity, which in an economic sense is executed by units. The various approaches and frameworks deal with these activities in different ways. In the routes of improvement discussed in this study activities are also treated in a specific way.

The first opportunity of improvement discussed (chapter 11), is the Common Comparable Package (CCP). In this method, the uniqueness of activities is stressed as a necessary prerequisite to enlarge comparability between countries. Basic in the method is that all units in the economic process are characterised by activities they perform. However, activities are not the carriers of information. Information on activities is supplied by units, including the providers of care. In the CCP study a list of activities¹⁾ is developed, for want of an adequately detailed and harmonised classification of activities. Because the starting point of this method is founded in the activities themselves, an enormous amount of information is necessary to be able to characterise all units sufficiently. This information was collected with the aid of a network of experts. In the CCP method, in the first instance units are clustered, by means of a bilateral comparison between two countries, in such a way that the contents of the total package of activities of both countries becomes (nearly) identical. In a comparison involving N countries, this procedure results in N-1

bilateral comparable packages (BCP's) if one country is used as a key country. For a common comparable package, valid for all countries, these various BCP's need to be uniformed to one CCP. This CCP appeared to be possible. Moreover, it was determined that this CCP was stable in size, i.e. additions of countries would not influence the contents, the size of the package of activities. To be able to use this method, however, a large amount of very detailed information is necessary, as well as a wide knowledge of the health (care) systems to be compared.

The advantage of the CCP is that international agreement is present on the method that is the foundation of the CCP: the use of activities as the criterion for selection on comparability. The detected disadvantages of this method relate to the laboriousness in terms of uniform packages of activities to be able to make comparisons, the amount of information needed and the lack of a (link to a) theoretical framework.

The System of Health Accounts (SHA) is discussed in chapter 12. In principle this system, developed by the OECD and supported by the EU, uses the same starting points as the CCP. However, because the level of detail and knowledge needed to make a success of the CCP, as well as the amount of work needed, is not manageable for an international organisation on an annual basis, another line of approach was chosen. In this approach, too, the activities performed and the units are the starting points. Units in the area of health care in the various countries are not comparable, activities are. In this method, however, the providers are not clustered until comparability is present, but activities are clustered into manageable uniform clusters, called functions. Since these functions are executed in flexible compositions by all units, ultimately they supply the basic comparability in the SHA. Various classifications have been developed in the SHA (classification of functions, classification of financing units and a classification of providers). Ideally using this method provides a tri-axial system, in which data can be presented with their dependent relations. Within the SHA, special attention is paid to the possibilities of linking these classifications to the classifications used in national accounts.

The advantages of the SHA relate to the presence of a theoretical foundation and strong embedding in the SNA. The disadvantages concern mainly the compromises following the composition of the classification of functions. Not enough justice is done to the purity of functions, while a provider characteristic (mode of production) is wrongfully mixed with function characteristics.

The Dutch System of Care Accounts (see chapter 13) – the successor to the cost and financing of health care statistics – is based on the SHA. In the utilisation of the structures of the SHA, various user targets are defined, for which different classifications of units are the starting points. In the process of starting and implementing, this new statistical system the aforementioned imperfections showed up. Improvements on the SHA have been carried through relating to the purity of the functions used. The incorrect addition in the SHA of production modes

to the functions instead of to the providers of care, is rectified. The Dutch System of Care Accounts presents data according to both the OECD classifications and the Dutch version of these classifications.

Lastly, both the classification of functions (ICHA-HC) and the classification of providers (ICHAS-HP) were expanded with functions of social care (SC) and actors in the social area. The classification of health care providers was rebuilt to a classification of health and social care providers (ICHA-HSP).

In a next chapter (chapter 14), other possibilities to enlarge comparability of health care expenditure data are discussed shortly.

One entrance could be the International Classification of Diseases (ICD) developed by the World Health Organisation (WHO). Although different interpretations in using this classification are possible, expenditure by ICD or ICD categories can offer an improvement in comparability. While expenditure is linked to providers, the problems relating to the identification of units is identical to the ones discussed in other methods. Summed over the totality of ICD codes this method supplies comparable data if the contents of the units to which the expenditure relates is identical.

For all other possibilities to be used (e.g. ESSPROS, European System of integrated Social Protection Statistics, DRG, Diagnosis Related Groups) the problems are no different. In each case, variables (like expenditure) need to be linked to other variables (like diseases, functions, etc.). Comparability of the results depends on the comparability of the units that are the starting point.

It can be concluded that the road taken with the Common Comparable Package developed in the Netherlands by Statistics Netherlands and built upon by the OECD with its System of Health Accounts provides a clear improvement in comparability. However, the problem of comparability, even that of health care expenditure, has not yet completely been solved. A lot of theoretical and practical work still needs to be done to create further improvements.

The efforts needed to implement a SHA should not be underestimated. The aim is that these efforts will yield profits in the long run, by means of better international comparability. Although within the EU social policy remains the prerogative of national governments, better comparability of health care data will prove advantageous to all member states individually as well as to the EU as a whole.

Part 3: Data: use and reconciliation

The third area (chapters 15 and 16) concerns the use of data and the reconciliation of the discussed frameworks in aid of comparability. Data only become relevant if they are used. The following questions can be formulated on this area:

10. Which data are used in a national and an international context? Do users realise that data are not always comparable, and if so, does this realisation influence conclusions based on the results? Do they realise that a certain quantity and quality of data is necessary for any research?
11. Is it possible to reconcile the data resulting from the various systems used in cross-country comparability of health care expenditure to reach an acceptable level of comparability? And what should be changed or added to reach a more optimal reconciliation?

In chapter 15, the use of data is the central theme, both in a national and in an international context. On many occasions, national use is made of data available in international databases. In a fair number of cases, users forget to mention how comparable the data are. International articles and econometric studies of health care also sometimes forget to remark on comparability of the source information. They do not seem to realise how this may influence the conclusions in particular. A very important item in using data for research is the quantity as well as the quality of data. A certain (high) level of quantity and quality are a precondition for valid research.

In chapter 16, possibilities are investigated to reconcile data on health care expenditure in existing systems. This test is executed with Dutch data, as these data were available in a detailed enough amount. Next, the consequences are investigated of the results of the national synthesis in satisfying the needs of cross-country comparable health care expenditure. Comparability is measured in relation to the international system of health accounts, adapted to the Dutch demands of functions. A problem, however, is the level of detail needed in the data based on the classifications used (especially the classification of activities, but also the classification of goods and services and the classification of functions). Based on national accounts data it is shown that depending on the starting points (production or consumption), and the criteria used, only a limited level of comparability could be achieved. In my opinion the level of comparability reached by using other systems than health accounts is not satisfactory.

Chapter 17 presents conclusions relating to innovations in the area of cross-country comparability of health care expenditure. It also looks at the use of data in international research and the use of data originating from various frameworks (existing and new ones) in the light of cross-country comparability.

From the analysis of the innovations, their possibilities to enlarge cross-country comparability and the use of data, the following remarkable observations can be deduced:

- Comparability can only be achieved by using the activities. Activities are a leading principle in every attempt to improve cross-country comparability of health care expenditure.

- Cross-country comparability can be improved further by developing a classification especially for this goal. The classification of functions (in a pure form) could be such a vehicle.
- In using data in international comparative research, more remarks on the level of (in)comparability of the data are needed. The influence of the lack of comparability on the results in particular should be specified more clearly. The use of data originating from existing systems, especially national accounts, offers only a limited level of usability. The main reason is that the link with the newly developed classification of functions can only be made with a large effort.

Recommendations

In the future the following improvements in comparability of health care data across countries could be reached (see chapter 18):

- *On activities.* Because activities play a crucial role in all frameworks (existing and new), it is important that a complete internationally agreed classification of activities is created, embedded in the classification of functions. A first initiative for such a classification of activities is presented in the Common Comparable Package (CCP) project. Completion of this list of activities and expansion with activities in social care as well as linking these activities with existing classifications is necessary.
- *On functions.* Because uniform functions (i.e. clusters of activities) are very important in the idea of the System of Health Accounts (SHA), the purity of these functions should be emphasised much more strongly. Mixing of parts of functions with each other and mixing with other aspects, like modes of production, should be excluded through international consultations.
- *On branch integration.* Because the dividing lines between the areas of health care and social care are fading, it is suggested that these two areas be integrated. Inclusion in the classification of providers as a separate cluster of actors should be contemplated. Expansion of the classification with social care functions should be researched and perfected.
- *On linking with national accounts.* It is obvious that the links with the internationally agreed system of national accounts should be made, if these new systems in specific areas can exist in their own right. How these links should and could be made (top-down from the aggregated systems of the SNA or bottom-up from the detailed information of e.g. the SHA) should be investigated further. In this study, first attempts in this area are presented in the form of connecting tables.
- *On the classifications of products.* The CPA/CPC (Statistical Classification of Products by Activity in the European Economic Community/Central Product Classification) could be usable for cross-country comparisons in health and social care if the classifications as they stand should be used to their fullest extent. However, to provide full comparability the classifications will have to be

deepened and broken down into more detail in the area of health and social care. Some starting points for this deeper level of detail can be found in the list of activities compiled for the CCP approach.

These methods will not result in complete comparability. Co-operation and co-ordination between all organisations concerned and their members will be necessary in order to realise further improvements.

Note:

- ¹⁾ In the annexes to Chapter 11 on the CCP the first results are presented on the links between the list of activities and the classification of functions of the OECD (ICHA-HC) as well as the classification of goods (CPA)

Abbreviations (used in the Summary)

BCP	Bilateral Comparable Packages
CBS	Centraal Bureau voor de Statistiek (Statistics Netherlands)
CCP	Common Comparable Package
CPA	Statistical Classification of Products by Activity in the European Economic Community
CPC	Central Product Classification
DRG	Diagnostic Related Group
EMU	European Monetary Union
ESA	European System of Accounts
ESSPROS	European System of integrated Social Protection Statistics
EU	European Union
HSP	Health Care and Social Care Providers
ICD	International Classification of Diseases
ICHA	International Classification for Health Accounts
ICHA-HC	International Classification for Health Accounts – Health Care Functional Classification
ICHA-HP	International Classification for Health Accounts – Health Care Provider Classification
FC	Functional Classifications
NA	National Accounts
OECD	Organisation for Economic Co-operation and Development
SA	Satellite Accounts
SC	Social Care Functions
SHA	System of Health Accounts
SNA	System of National Accounts
SP	Social Care Providers
WHO	World Health Organisation

Contents

Acknowledgement	V
Samenvatting	VII
Afkortingen (gebruikt in de Samenvatting)	XIX
Summary	XXI
Abbreviations (used in the Summary)	XXXI
1. Introduction	1
1.1 Background	1
1.1.1 The national level	1
1.1.2 The international level	3
1.2 International comparability	4
1.3 Data availability and comparability	5
1.4 Thesis and questions	10
1.5 Sequence of the chapters	12
2. Terms and definitions used in the title	15
2.0 Questions	15
2.1 Introduction	15
2.2 Comparability and related terms	15
2.2.1 Metadata and comparability	19
2.3 Health and health care	21
2.3.1 Health statistics	23
2.3.2 Health care statistics	26
2.3.3 Production boundary in relation to health care	27
2.4 Cost and expenditure	29
2.5 Final remarks	31
<i>Part 1: Description and analysis of existing economic frameworks</i>	
3. System of National Accounts/European System of Accounts (SNA/ESA): relations of aggregate economic activity	35
3.0 Questions	35
3.1 Introduction	35
3.2 SNA/ESA: background	35
3.3 A System of National Accounts	37
3.3.1 The SNA: criteria and starting points	38
3.3.2 The development of the SNA	39
3.3.3 Functional classifications and functional approach	40
3.4 The European System of Accounts	42

3.4.1	The ESA: criteria and starting points	42
3.4.2	The development of the ESA	43
3.5	SNA, ESA and comparability	44
3.5.1	Concepts and definitions	44
3.5.2	Valuation and units	45
3.5.3	Classifications	45
3.5.3.1	Development of classifications in the SNA and the ESA	45
3.5.3.2	SNA and classifications	46
3.5.3.3	ESA and classifications	46
3.6	SNA/ESA and health care statistics	47
3.6.1	Introduction	47
3.6.2	Production	47
3.6.3	Consumption, expenditures and financing	49
3.7	Comparability and conclusions	51
3.7.1	Comparability and SNA/ESA	51
3.7.2	Comparability and standards	53
3.8	Final remarks	54
4.	Satellite accounts	57
4.0	Questions	57
4.1	Introduction	57
4.2	The history of satellite accounts	58
4.3	Scope of satellite accounts	60
4.4	Satellite accounts: definitions and descriptions	62
4.4.1	Satellite accounts: general description, features and characteristics	62
4.4.2	Satellite accounts: framework and definitions	63
4.4.3	Description of satellite accounts on health care	64
4.5	Contents, scope and construction of satellite accounts on health care .	65
4.5.1	Introduction	65
4.5.2	Contents, scope and construction	65
4.6	Satellite accounts on health care and comparability	67
4.7	Final remarks	70
5.	Functional classifications	73
5.0	Questions	73
5.1	Introduction	73
5.2	Description of functional classifications	73
5.3	The history of functional classifications	74
5.4	Objectives of functional classifications	75
5.5	Functional classifications on health care	76
5.6	Functional classifications on health care and comparability	78

5.6.1	Advantages or favourable conditions	79
5.6.2	Disadvantages or unfavourable conditions	80
5.7	Conclusions on functional classifications as an instrument for comparability	81
5.8	Final remarks	81
6.	Health accounts	85
6.0	Questions	85
6.1	Introduction	85
6.2	History of health accounts	86
6.3	Aims and objectives of health accounts	88
6.4	Starting points and criteria	88
6.4.1	Criteria	89
6.4.2	Features and characteristics	89
6.5	Health accounts framework: results	90
6.6	Health accounts and comparability	91
6.7	Final remarks	94
7.	OECD health database	97
7.0	Questions	97
7.1	Introduction	97
7.2	History of the health database	97
7.3	Purposes	99
7.4	Approaches to comparability	100
7.5	OECD health database and comparability	101
7.5.1	Introduction	101
7.5.2	Multiple sources	101
7.5.3	Data comparability	102
7.5.4	Metadata and converging concepts	103
7.5.5	Contents of items and comparability	103
7.6	Final remarks	104
8.	The statistical process	107
8.0	Questions	107
8.1	Introduction	107
8.2	Input process	108
8.3	Throughput process	111
8.4	Output process	113
8.5	Final remarks	114
9.	Comparability of existing data: a practical example	117
9.0	Questions	117
9.1	Introduction	117

9.2	Data sources on health and social care	118
9.3	Health and social care shares in SNA/ESA aggregates	122
9.4	Health care data in national health accounts of the USA	123
9.5	Health care data in the OECD health database	126
9.6	Time series analysis	128
9.7	Final remarks	132
10.	Existing frameworks: conclusions	135
10.0	Questions	135
10.1	Introduction	135
10.2	Usability and comparability	135
10.2.1	Frameworks and relations	136
10.2.2	Frameworks and valuation	137
10.2.3	Frameworks and units	138
10.2.4	Frameworks and classifications	140
10.3	Frameworks and the differences in production (value) and consumption (expenditure)	141
10.4	Conclusions based on results of the frameworks	142
10.4.1	Conclusions on SNA and ESA	142
10.4.2	Conclusions on satellite accounts	143
10.4.3	Conclusions on health accounts	144
10.4.4	Conclusions on the OECD health database	144
10.4.5	Conclusions on functional classifications and the statistical process	144
10.4.6	Conclusions on data collections	145
10.5	Final remarks	146

Part 2: Innovations concerning cross-country comparability

11.	Common Comparable Package (CCP)	151
11.0	Questions	151
11.1	Introduction	151
11.2	General approach	152
11.3	Process description and the use of metadata	154
11.4	Methods and instruments	155
11.4.1	Classifications	155
11.4.1.1	'Classification of services' applied in the CCP approach	155
11.4.1.2	The list of activities in health care and international classifications	156
11.5	Method and model	158
11.6	The Bilateral Comparable Package (BCP) process	160
11.7	The Common Comparable Package (CCP) process	161

11.7.1	Introduction	161
11.7.2	From Input to CCP: additions and subtractions	161
11.7.3	Results: constructed input expenditure file, BCP and CCP ..	164
11.7.3.1	Constructed input expenditure file and BCP compared	165
11.7.3.2	Expenditure data in the OECD health database and the CCP approach compared	165
11.8	CCP cure and care	166
11.9	The CCP approach and comparability	169
11.10	Final remarks	170
12.	System of Health Accounts (SHA)	173
12.0	Questions	173
12.1	Introduction	173
12.2	History	174
12.3	Criteria for the SHA	174
12.4	SHA classification system	175
12.4.1	Classification of functions	176
12.4.2	Classification of providers	177
12.4.3	Classification of financing units	179
12.5	SHA and SNA: health care boundary, production boundary, subsidies and expenditure	179
12.5.1	Health care boundary definition: SHA compared with SNA	180
12.5.2	Household production: SNA and SHA compared	180
12.5.3	Subsidies and transfers	181
12.5.4	Concept of expenditure	181
12.6	SHA and satellite accounts	182
12.6.1	Differences between SHA and satellite accounts	182
12.6.2	Similarities between SHA and satellite accounts	183
12.7	Problems in the SHA in relation to comparability	183
12.7.1	Discussion of the classifications	183
12.7.1.1	Terms used in the classification	184
12.7.1.2	Characteristics linked to the classification	184
12.7.1.3	Internal consistency	185
12.7.2	Concept of expenditure and comparability	187
12.7.3	Treatment of transfers and subsidies	188
12.8	SHA in relation to the OECD health database	189
12.9	Final remarks	190
13.	Implementation of the SHA in the Netherlands and cross-country comparability improvement	193
13.0	Questions	193
13.1	Introduction	193

13.2	Background: reasons and objectives	194
13.3	Methodology	195
13.4	National results	198
13.5	Cross-country 'comparable' results using connecting tables	202
13.5.1	User needs and connecting tables	202
13.5.2	Connecting tables and comparability of results	203
13.6	Expansion of the classification of providers and the classification of functions	207
13.6.1	Classification of providers	207
13.6.2	Classification of functions	208
13.7	Final remarks	211
14.	Other 'new' developments	215
14.0	Questions	215
14.1	Introduction	215
14.2	Other 'new' developments and comparability	215
14.2.1	Developments relating to 'expenditure'	215
14.2.2	Developments relating to 'health' or 'health care'	218
14.3	Final remarks	219

Part 3: Data: use and reconciliation

15.	Users and uses of data	223
15.0	Questions	223
15.1	Introduction	223
15.2	International data used in the Netherlands	223
15.3	Data use and comparability	224
15.3.1	Comparability of health care expenditure in policy documents	225
15.3.2	Comparability of health care expenditure used in international reports	225
15.4	Health care expenditure comparability and analysis	227
15.4.1	Health care expenditure and cross-country comparability ..	227
15.4.2	Health systems performance discussions and cross-country comparability	230
15.4.2.1	Introduction	230
15.4.2.2	Attainment and performance: global list	231
15.4.2.3	Methods and data used	233
15.4.2.3.1	Discussion of methods used in the World Health Report 2000	233
15.4.2.3.2	Data availability and data use in the World Health Report 2000	236
15.5	Final remarks	239

16. Reconciliation	243
16.0 Questions	243
16.1 Introduction	243
16.2 Starting points	244
16.3 National Accounts and reconciliation	246
16.3.1 Activities and related boundaries	246
16.3.2 Terminology: production value and expenditure	247
16.3.3 Valuation of production	247
16.3.4 Reconciling National Accounts and Care Accounts	248
16.3.4.1 Health care production	248
16.3.4.2 Health care production and final consumption ...	251
16.3.4.3 Health care production and expenditure	252
16.4 The use of the statistical classification of products by activity in the European Economic Community (CPA)	253
16.5 Functional classifications and reconciliation	255
16.6 National data systems and cross-country comparability	256
16.7 Final remarks	258
 17. Conclusions on new approaches and use of data	261
17.0 Questions	261
17.1 Introduction	261
17.2 Conclusions on the CCP approach and the SHA method	262
17.2.1 Activities versus functions	262
17.2.2 Theoretical links to the System of National Accounts, SNA, and the European System of Accounts, ESA	262
17.2.3 Other aspects	264
17.3 Conclusions from 'other' developments	264
17.4 Conclusions from reconciliation of methods	265
17.5 Conclusions from users and uses of data concerning comparability	267
17.6 Final remarks	268
 18. Final conclusions and recommendations	271
18.1 Introduction	271
18.2 Final conclusions	271
18.3 Recommendations	273
 Annexes	275
References	303
Curriculum vitae	311
Abbreviations	313
Descriptions of most important terms used in this study	317

Figures		
Figure 2.1	Model used in Public Health Status and Forecasts	24
Figure 2.2	Overview of relations regarding health	25
Figure 2.3	Production boundary	28
Figure 3.1	Circular flows including government	36
Figure 3.2	Theoretical approach to comparability	54
Figure 5.1	Comparability and functional classifications	78
Figure 8.1	Statistical process	108
Figure 10.1	Relations between production and consumption	136
Figure 10.2	Valuation and gross turnover	138
Figure 10.3	SNA, Satellite Accounts and Health Accounts: production compared	139
Figure 10.4	Co-ordination, detail and comparability	146
Figure 11.1	Process of the construction of bilateral comparable packages and the common comparable package	153
Figure 11.2	Method and model	159
Figure 11.3	Cure and care providers versus cure and care functions	168
Figure 13.1	Construction process health and social care accounts	196
Figure 13.2	'Users' of the system and connecting tables	202
Figure 13.3	Health and social care system	209
Figure 13.4	Health and social care functions	210
Figure 14.1	Household budget survey and health care expenditure	216
Figure 14.2	Social protection and health care expenditure	217
Figure 16.1	Reconciling existing systems and cross-country comparability: main problem areas	245

Graphs		
Graph 1.1	Average Length of Stay (ALOS) in in-patient care institutions, 1995	6
Graph 1.2	Average Length of Stay (ALOS) in in-patient care and acute care institutions, 1995	7
Graph 1.3	Expenditure on in-patient care institutions as a percentage of GDP, 1995	7
Graph 1.4	Expenditure per nursing day in in-patient care institutions in US \$, 1995	8
Graph 1.5	Expenditure per nursing day in acute care institutions in US \$, 1995	9
Graph 9.1	Health and social care as a percentage of GDP in EU Member States by source, 1997	119
Graph 9.2	Health and social care as a percentage of GDP (including and excluding health and social care) in EU Member States, 1997 ..	121
Graph 9.3	Personal health care percentage of GRP in the USA, 1998	124

Graph 9.4	Personal health care percentage of GRP and GSP in the USA, 1998	125
Graph 9.5	Health care expenditure of OECD and EU member states, percentage of GDP, 1997	127
Graph 9.6	Health care expenditure percentage of GDP in the Netherlands, 1953–1999	128
Graph 9.7	Real growth per capita of health care expenditure in the Netherlands, 1953–1999	129
Graph 9.8	Real growth per capita of health care expenditure in two OECD countries (France and the United Kingdom), 1972–1998 .	130
Graph 9.9	Health as percentage of GDP compared for four EU Member States (the Netherlands and Belgium versus France and the United Kingdom), 1972–1998	131
Graph 13.1	Expenditure by function: OECD-HC and Dutch-‘HC’ classification of functions compared, 1998	206
Graph 15.1	Attainment & Performance and constituent factors for the European Union	232

Tables

Table 3.1	Overview of classifications used in SNA 1993 and ESA 1995 ...	46
Table 3.2	ISIC and NACE relevant to health in SNA/ESA	48
Table 5.1	Functional classifications on health care	77
Table 9.1	Health and social care as a percentage of GDP in EU Member States by source, 1997	118
Table 9.2	Share of NACE class N in various National Accounts aggregates, 1997	122
Table 11.1	Non-BCP provisions	162
Table 11.2	Added calculations in constructing the various Blocks	163
Table 11.3	Additions in constructing the CCP	163
Table 11.4	Subtractions in constructing the CCP	164
Table 11.5	Constructed input expenditure file, BCP and CCP in ECU, 1995	164
Table 11.6	Total health care expenditure in CCP approach and OECD health database as a percentage of GDP, 1995	166
Table 11.7	The CCP of Block I: In-patient care and Block II: Outpatient care divided into cure and care, 1995	167
Table 12.1	ICHA-HC Classification	176
Table 12.2	ICHA-HP Classification	177
Table 12.3	ICHA-HF Classification of sources of funding	179
Table 12.4	The OECD classification of functions and a ‘pure’ classification of functions	186
Table 13.1	Classification of functions, Dutch variation	198

Table 13.2	Expenditure by main user perspective, million euro, 1998–2001*	199
Table 13.3	Expenditure care accounts by source of funding, million euro, 1998	200
Table 13.4	Expenditure care accounts by function, million euro, 1998	201
Table 13.5	Expenditure care accounts by source of funding, according to OECD ICHA-HF classification, million euro, 1998	204
Table 13.6	Expenditure care accounts by function, according to OECD ICHA-HC classification, million euro, 1998	205
Table 15.1	Selection of 1st and 2nd Generation Research	227
Table 15.2	Data: gathered and imputed, 1997	238
Table 16.1	Classification rules and the boundary of health care	246
Table 16.2	National Accounts data: total economy and branches, million euro, 1998	249
Table 16.3	National Accounts and Care Accounts compared, million euro, 1998	250
Table 16.4	SNA production value and Care Accounts receipts by major categories of differences, million euro, 1998	251
Table 16.5	Production and final consumption of health care: SNA and Care Accounts, million euro, 1998	252
Table 16.6	Calculation of production and expenditure, million euro, 1998	253
Table 16.7	Health care by product and Care Accounts expenditure, million euro, 1998	254
Table 16.8	Health care by function and Care Accounts expenditure, million euro, 1998	255
Table 17.1	Coverage of SNA aggregates in relation to the Care Accounts, 1998	265
 Annexes		
Annex 11.1	List of Activities in health care in the CCP approach	275
Annex 11.2	Adapted List of Activities linked to functions in the EUCOMP project	278
Annex 11.3	CPA 1996 and NACE rev. 1 cross-classified	280
Annex 11.4	Adapted List of Activities linked to functions in the EUCOMP project and the relevant CPA 1996 code(s)	282
Annex 11.5	Adapted List of Activities linked to functions in the EUCOMP project and the relevant NACE rev.1 code(s)	284
Annex 13.1	Health and Social Care Provider Classifications	286
Annex 13.1A	Providers in the Care Accounts cross-classified by OECD HSP extended code, OECD ICHA-HP code, SIC NL code, ISIC/NACE code and Care Statement chapter	286

Annex 13.1B	Providers in the Care Accounts by OECD ICHA-HP code cross-classified by OECD HSP extended code and Care Accounts code	288
Annex 13.1C	Providers in the Care Accounts by ISIC/NACE code cross-classified by Care Accounts code and OECD HSP extended code	291
Annex 13.2	Definitions of the Functions of Health Care in the Netherlands	293
Annex 13.3	Definitions of the Functions of Social Care in the Netherlands .	297
Annex 13.4	Activities linked to ICHA-HC of the OECD and the Dutch HC Classification	299
Annex 16.1	Relation between SNA by ISIC/NACE and Care Accounts categories, 1998	300

1. Introduction

1.1 Background

Not only have technical possibilities in our present society facilitated an enormous increase in the amount of information becoming available, it has also made it easier to access this information, for example by e-mail and the Internet. This is also true of statistical information; indeed the relative importance of statistical information within the overall demand for information is growing. Users want it faster, they want more and they want it in greater detail. They use it to negotiate agreements (e.g. employers vs. employees), for political decision making, and in the evaluation of political processes and their consequences. Statistical information includes not just financial or monetary data, but all the information used in the processes described. There is also a need and demand for information in general and statistical information in particular at an international level; and here too it is extensive and it is increasing. The European unification, the introduction of the monetary union and the globalisation of the world economy have prompted a growing need for information which is reliable, comparable and harmonised for the European Union and other, broader collaborative networks. Globalisation demands more comparable data in a technical sense. This need can be strengthened by political processes like the European integration to ensure transparency. The priority given to comparability is determined by these two objectives.

The important policy area of public health and health care is also affected by these comprehensive and far-reaching developments. The need for correct and adequate information is felt on national as well as international levels.

In the national context, like their counterparts in the business sector, managers of health care institutions need accurate and reliable statistical information. Supervising bodies in every country (like the Commission on Tariffs in Health Care, CTG, in the Netherlands), need to underpin control and budget decisions with detailed information. Ministerial departments need timely information both to respond to legislative controllers (Parliament) and to requests from political parties to substantiate political decisions. Internationally comparative data are especially needed by the ministerial bodies and international judicial organisations such as the controlling bodies of the health insurance funds to decide on the legitimacy of rules and regulations in preparation, in relation to European legislation, or claims of nationally insured persons for provisions received abroad.

1.1.1 The national level

Health care is an important issue in society, the political arena and the economy. The constant growth of the health care industry has exerted an enormous pressure on

available resources. If we are to solve the present problems, anticipate future challenges and create an adequate policy in the field of public health and health care, a system needs to be developed that can provide detailed and comprehensive information. Information is needed on the whole system, from the input (manpower, material etc.), through the process (activities in the various institutions and performance of self-employed) to the output and the outcome. Output can be determined as the products defined in terms such as Diagnosis Treatment Combinations (DTC), Diagnosis Related groups (DRG), etc. The outcome of the health care system can be described as the gain in health for specific groups in the population, exemplified by a reduction in potential life years lost (PYLL), a gain in disability adjusted life years (DALY) or health adjusted life expectancy (HALE). It goes without saying that information on prices and quantities delivered is also badly needed, not only physical data but also data on cost/expenditure, receipts and the financing of health care. Nations are increasingly aware of the importance of the financial dimensions of health care systems. Quality information is necessary to understand the functioning of any health system, to realise the need for reform and to measure the impact of reforms. Health care financing is an important indicator of a system's status and change and therefore plays an increasingly important role in the policy agenda of providers and managers of health care services.

Health care is either publicly funded, from taxes or insurance schemes, or privately funded, through private insurance schemes and out-of-pocket payments or financed by a combination of public and private funding. Consequently a balance between an adequate level of health care and its related expenditure is a policy target, just like any other public sector.

Across Europe, national health care systems are currently confronted with enormous challenges. These are mainly caused by ageing populations which result in more people in the categories needing care, but factors such as the increasing demand for care and continuous technological progress are also playing a part. Expenditures will rise. Social solidarity will be under pressure, if indeed it is not already. This is important because resources are scarce and must be used rationally. These developments contribute to an increasing attention for the question of expenditure. Health care expenditure can be used as a tool in economic policy.

In times of low and slowing economic growth (e.g. the 1980s), attention is even more focused on areas like health care expenditure and financing. Economic problems prompted policymakers to look over the fence: were governments in other countries doing better or worse? What can be learned from national policies and systems there? Governments and politicians need information in general and statistical information in particular to be able to answer these and similar questions.

In many countries private funding of health care appears to be increasing, and this is one of the arguments that has contributed directly to a larger interest from private insurance companies and consumer organisations in developments in the field of health care.

In the Netherlands, too, attention has shifted from government intervention to more market-oriented influences in health care (the so-called Dekker proposals¹⁾), based on the notion that private mechanisms could better clear markets than government intervention could. At the same time the supply-controlled mechanisms were to be replaced by demand-induced ones.

1.1.2 *The international level*

In an international perspective there is a wide range of actors. For health and health care, the WHO (World Health Organisation, a United Nations organisation), the OECD (Organisation for Economic Co-operation and Development), some other United Nations bodies and of course the European Union and its bodies are of particular importance.

The WHO is an old and respected actor in the international health arena. It originated from the Health Department of the League of Nations²⁾, whose "... main objective [was] to prevent contagious diseases on a global level by means of internationally accepted rules on control and decontamination." (Querido, 1967, page 128). The WHO is mainly concerned with the health and the development in health in the populations of its member countries. Health and health status are important topics in this domain. According to its recently published World Health Report 2000 (Overview, page XI), all governments have three objectives: improving health, a fair allocation of financing, and performance of the system. To measure how countries are doing with respect to these objectives, a rich body of information on health and health care systems in all countries is needed.

The OEEC (Organisation for European Economic Co-operation) was established as a direct consequence of the need for collaboration in all economic areas after the end of World War II (e.g. the introduction and monitoring of the Marshall aid in Europe). In 1961 the OEEC was succeeded by the OECD. The OECD became active in the health and health care field in the mid-seventies. The main area of its concern is economic policy in a broad sense. The most important topics of the OECD relate to health care data in an economic sense: expenditures, human resources and output. The UN has an even broader scope than the OECD. For the health care branch the importance of the UN is largely in the impressive number of economic classifications: the ISIC (International Standard Industrial Classification), COICOP (Classification of Individual Consumption by Purpose), COFOG (Classification of Functions of Government), COPP (Classification of Outlays of Producers by Purpose), COPNI (Classification of Purposes of Non-Profit Institutions Serving Households) and other classifications play an important role. Of course the development of international frameworks like the SNA (System of National Accounts) should also be mentioned in this respect. Various UN institutions have developed classifications that eventually influence rules to be implemented in health accounting.

Within the European Union, Eurostat and the Directorate SANCO (Health and Consumer Protection) are the most important international players; Eurostat in the statistical area, SANCO on a political level. Eurostat has a duty to supply the Commission, its entities and Member states with statistical information on all areas relevant to EU policy.

The Treaty of Rome (1957) introduced a common philosophy for the economic and social areas. In the Treaties of Maastricht 1992 and Amsterdam 1996 the need for co-operation in the areas of health policy and priorities in health care are mentioned explicitly. Sections 36(o) and 129 of the Maastricht Treaty supply the basis for community policy in the public health area. The first Health Monitoring Programme (1997–2001) ³⁾ has been one of the means to this end. On the other hand, the areas of social care – of which health care is an important sub-area – are the sole responsibility of the Member states. The absence of a direct EU mandate for steering the areas of health and social care, however, does not mean that the EU does not exert any influence on existing systems. Especially the EU Treaties, comprising the freedom of circulation of persons and goods (and services), are an inherent driver of harmonisation of health care systems. In the shorter or longer term the existence of the EU will force the differences between health systems to become smaller (see e.g. SCP 2000, page 264). At a statistical level this lack of power has consequences for the products the Member states supply to the Union. By no means countries can be forced to change their own system.

International health care focuses on more or less the same topics as the national policy areas: equity, quality of the system, and the level of expenditure, especially the rise in expenditure.

1.2 *International comparability*

All EU Member states produce health care statistics – some have done so for a very long time – based on national traditions, laws and regulations and primarily intended for use in a national context to serve national needs. These statistics are not comparable with each other because of differences in the organisation of health care (the infrastructure), in financial systems, in the providers of medical services and in the boundaries with other sectors (like social security, government, etc.).

All inter- and supranational organisations face the same problems in this respect: the lack of comparability in the available data. The OECD and WHO have done a lot of work in the area of international comparison, formulating regulations on how to convert national data to international data. The actual transformation of the data at the national level is in the hands of the Member states. It is a very labour intensive process and not always well co-ordinated. Co-ordination requires more structured consultations between individual countries and between countries and international organisations. The EU is trying to expand international comparability

by advising standardisation at the micro-level and by improving data exchange possibilities.

International comparisons are only as good as the data on which they are based. By the very nature of administrative monitoring requirements, in all countries, data are primarily collected for national purposes. Differences between countries remain: different administrative systems and data sources, different classifications and definitions of variables, and different terminology and collection methods. The boundaries of the health care system used at the national level place restrictions on the cross national comparisons. When sub-categories of the systems are to be compared, the boundary problem becomes even more important, because definitions of sub-categories of health care may differ even more than overall boundaries used.

Different accounting systems add to the problems: for example use of the System of National Accounts (SNA), the European System of Integrated Social Protection Statistics (ESSPROS), the National Health Accounts, (NHA). Furthermore, health reforms may exacerbate existing problems (e.g. shifting areas between branches).

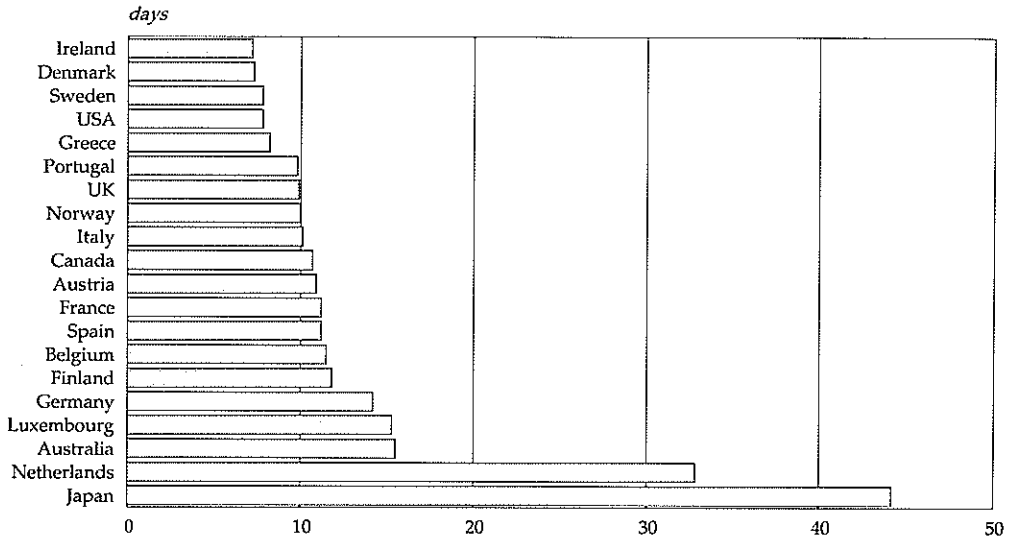
Although much work still has to be done, it is now accepted that more co-operation and standardisation is needed between countries to achieve a higher level of, and thus more useful, comparability in statistics produced by individual countries.

The OECD Health Database would seem to be a good starting point for this work. The OECD has a longstanding knowledge in the area of health care and the problems in making health care expenditures comparable. Moreover, all EU Member states – sometimes stimulated by some pressure of Eurostat – supply data to the OECD.

1.3 Data availability and comparability

The existence of data collections of international organisations – like the OECD health database, that started in 1982 – made it easier to make comparisons between countries. For the first time attempts were made to collect more or less comprehensive sets of data on the health care area with the aim of presenting data for cross-country comparisons that would allow users to create an international ranking, as well as giving them the opportunity to learn from each other's results. The database was filled with data of member countries. Analysis of the material by the OECD led to the conclusion that the data as such were very valuable but not comparable at all levels. At detailed levels of the health care systems in particular, data incomparability became apparent. It had already become clear at the start of the project that the most prominent causes of this incomparability were the various starting points of the data suppliers with regard to boundaries, contents and definitions of the health care area.

1.1 Average Length of Stay (ALOS) in in-patient care institutions, 1995



Source: OECD Health Data 2000.

An example of this incomparability is presented ⁴⁾ in the first graph (graph 1.1) containing data on the average length of stay (ALOS) in in-patient care institutions for 1995.

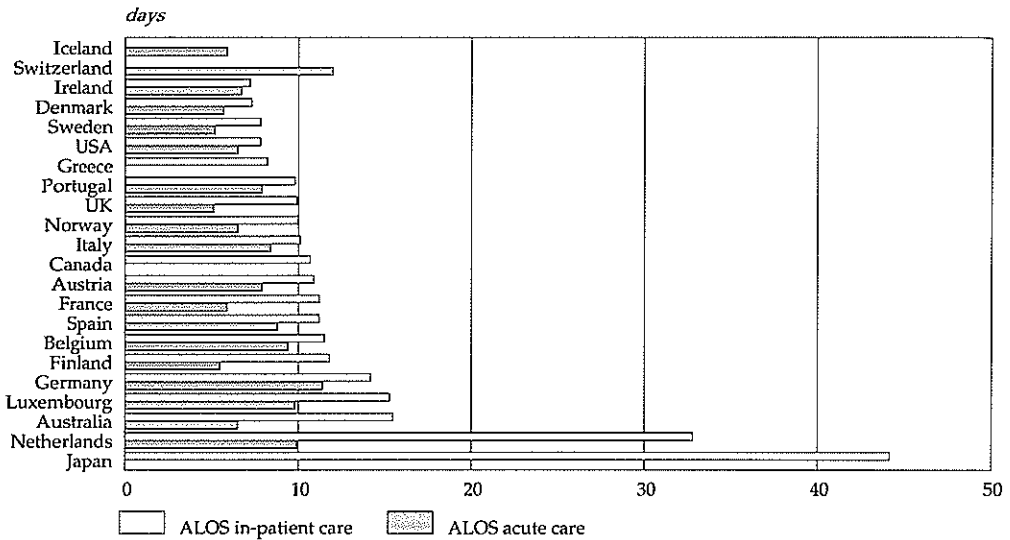
The ALOS varies from 7.2 days in Ireland to 44.2 days in Japan. Is it justified to conclude that the ALOS of in-patient care in Japan may not be compared with that in Ireland because the difference is so large? Or do the figures allow to conclude that the ALOS of Portugal, the United Kingdom, Norway and Italy are comparable, because the differences are so small?

The degree of incomparability is less pronounced for the ALOS of acute care, but the ALOS still varies from 5.1 days in the United Kingdom to 12.0 days in Switzerland. The impression that thus emerges is that the contents of the categories of institutions for which the ALOS is calculated are not the same (aside from the way the ALOS is calculated). To justify these conclusions, however, one would have to be sure that the contents of the term, the calculation or estimation procedures and measurement of the basic data used in the calculations are more or less identical.

Combining the two series on ALOS, for in-patient care and for acute care, this incomparability becomes more pronounced (more likely). This picture is illustrated in graph 1.2.

Some countries (e.g. Iceland and Switzerland) have no data on ALOS in in-patient care institutions, but do have data on the ALOS in acute care. On the other hand there are countries (e.g. Greece, Canada and Japan) that present data on the ALOS in in-patient care institutions but apparently do not have any data on the ALOS in acute care.

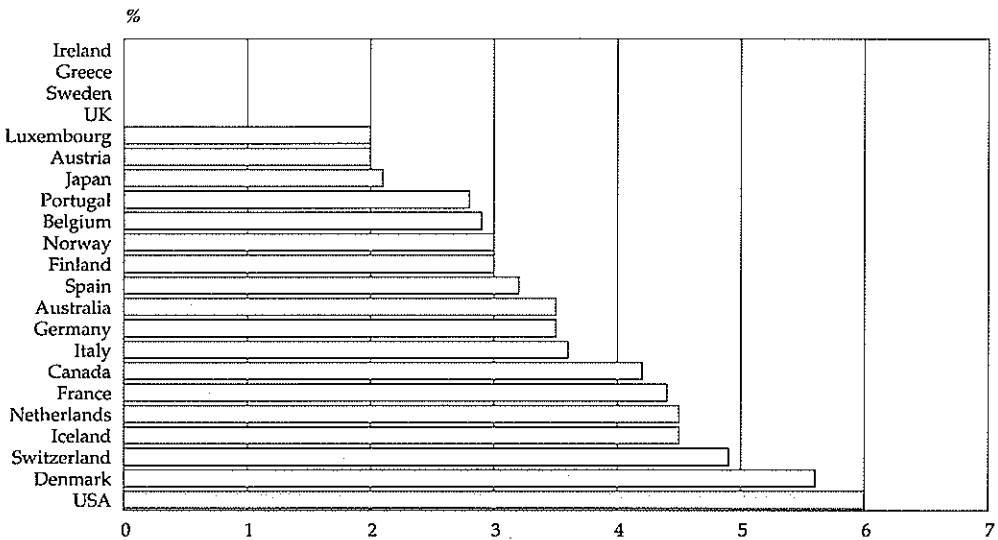
1.2 Average Length of Stay (ALOS) in in-patient care and acute care institutions, 1995



Source: OECD Health Data 2000.

There appears to be a very limited difference in ALOS of acute care and in-patient care in Ireland. On the other hand, the enormous difference in ALOS in the Netherlands between these two categories is at least remarkable.

1.3 Expenditure on in-patient care institutions as a percentage of GDP, 1995



Source: OECD Health Data 2000.

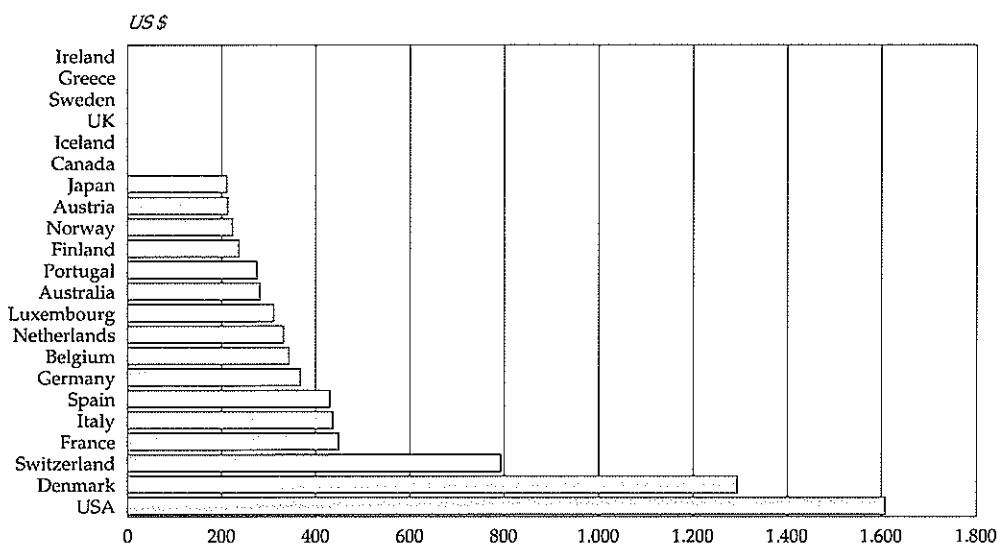
Another fact follows from the inter-country comparison of the ALOS for in-patient care with that for acute care. A higher in-patient care ALOS (compared with a certain country) does not necessarily correspond with a higher acute care ALOS compared with the same country (see e.g. Belgium and Finland, or Germany, Luxembourg and Australia). Belgium has almost the same level of ALOS for in-patient care compared with Finland but a higher ALOS for acute care.

The financial data of the countries presented in the OECD database reveal a similar lack of comparability, although less apparent even in the more detailed areas, like hospital care and in-patient care in general. In graph 1.3 the data on the expenditures on in-patient care (as a percentage of gross domestic product, GDP) are presented as an example.

The amounts spent on in-patient care as a percentage of GDP vary from 2.0 percent in Luxembourg to 6.0 percent in the USA, but are the boundaries of the system of health care the same? And do the calculation and classification schemes differ (i.e. are they based on the System of National Accounts, or on a national health accounts approach, satellite accounts or any other health accounting process, or maybe any variation in these frameworks and processes)?

However, if the data on expenditures in these detailed areas are linked to the uses of the health care system, say to the number of nursing days in in-patient care and the number of nursing days in the acute care area, then the same differences as those illustrated in the ALOS graphs become apparent. In the two graphs (graphs 1.4 and

1.4 Expenditure per nursing day in in-patient care institutions in US \$, 1995



Source: OECD Health Data 2000.

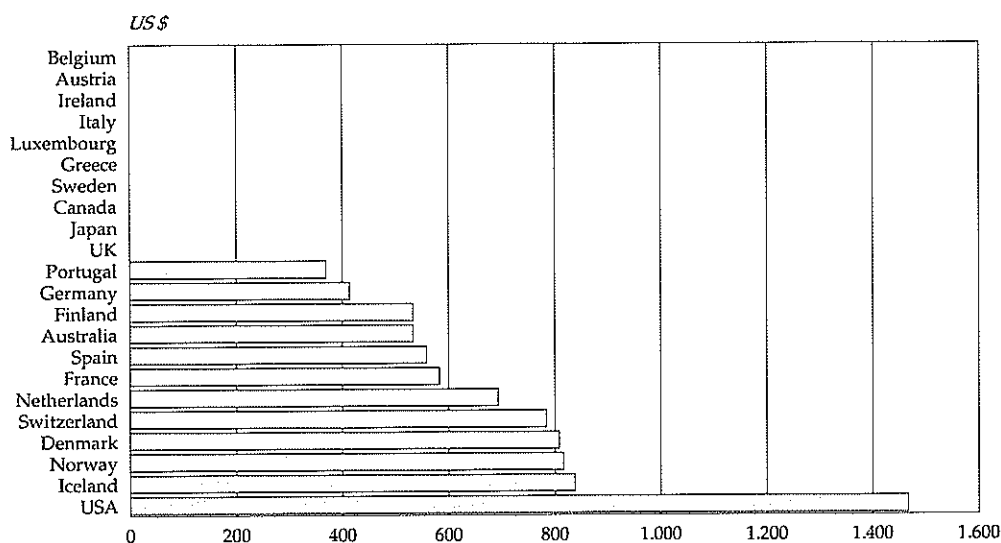
1.5), expenditures are linked to the number of in-patient days for the two areas: in-patient care and acute care.

No data are available on in-patient expenditures or nursing days for six of the 22 presented countries. Expenditure per nursing day for in-patient care was over \$ 600 in only three countries in 1995. If these data are comparable this would mean that in the USA in-patient care nursing days are twice as expensive as in Switzerland and almost eight times as expensive as in Japan or Austria.

Is it conceivable that the expenditures per in-patient day as presented in this graph are based on the same boundaries and contents of the area for in-patient care?

For acute care (graph 1.5) the same differences, although less pronounced, are visible.

1.5 Expenditure per nursing day in acute care institutions in US \$, 1995



Source: OECD Health Data 2000.

For in-patient care data, six countries had no available financial or nursing day data; in the area of acute care ten countries lacked data on either expenditures or acute care days. In the remaining twelve countries (excluding the USA) the average expenditures per acute care day vary from around \$400 to around \$800. Only in the USA does the amount spent per acute care day really deviate from the rest, at more than \$1,400, which is around 4.1 percent of (a much higher) GDP. The USA spends

around 14 percent of its GDP on total health care, which is much more than any other country in the world.

It is hardly conceivable that the definitions and boundaries of the area of acute care are the same in the USA as in Germany or Finland, for example, even though the relative price level of health care goods and services in the USA can be much higher compared to other countries.

Given that these differences are revealed in the OECD health database, and given that the data are supplied by national (statistical) organisations, it is no wonder that expenditure/cost data combined with disease data also provide data that are not comparable. In his thesis, Polder correctly concluded that "... the scope of COI-studies [Cost of Illness] barely extends beyond national levels." (Polder 2001, page 109). Furthermore, his study concludes that "... cross-national comparison should start with health care systems ..." (Polder 2001, page 109).

The apparent incomparability in the data shown above may be caused by numerous factors (e.g. different definitions and classifications, interpretation of rules, cultural differences). Comparability of the results can only be achieved if the contents of the sets of variables to be compared are harmonised or unified. Polder (Polder 2001, page 93, 94) distinguishes six explanatory factors ⁵⁾ for cross-national differences, which are not ranked arbitrarily. The two most important factors are health care supply and health care expenditures. These two factors are the focus of the study.

1.4 *Thesis and questions*

The present study focuses on the problems of international comparability of expenditure in the area of health care. The hypothesis and the starting point for this study is:

'The expenditure on health care of a selection of countries as presented in various, more or less integrated, existing databases is not comparable. However, there are possibilities to enlarge the comparability of these data.'

This complex and all encompassing issue is approached in three parts.

The first part (chapters 3 to 10) looks into a description of existing economic frameworks and the analysis of the usefulness of these systems for cross-country comparability of health care expenditure. Existing systems in this context are frameworks developed since the 1950s and broadly applied in many countries.

The second part (chapters 11 to 14) concentrates on the description and analysis of innovations aiming to enlarge comparability. It investigates the possibilities to increase the comparability in data on the health care expenditures, using new, recently developed methods and statistics. New methods in this study refer to

methodologies and systems although existing (since five to ten years) are still in a developing stage and not yet broadly applied.

The third part (chapters 15 and 16) is concerned with the use of data and the reconciliation of the existing frameworks with the innovations in relation to possibilities to enlarge comparability.

To be more specific: the study does not include topics concerning the inputs in the health care system (e.g. health care employment and material), nor the output of the health care institutions (products and services rendered), nor the outcomes ⁶⁾ in terms of a reduction of life years lost or disability free life years, nor health or health status.

The following questions can be raised with respect to the first area, the analysis of existing systems and their usability for cross-country comparability of health care expenditure:

1. Is comparability an issue in existing international frameworks like the System of National Accounts (SNA) and the European System of Accounts (ESA)?
2. Are the starting points and criteria mentioned in for example the SNA and ESA hampering or fostering comparability, especially comparability of health care data?
3. Does the development of new versions of the SNA and ESA lead to improvements in comparability in general terms, and in health care terms in particular?
4. Does the level of comparability increase when more detailed systems like satellite accounts or (functional or national) accounts are used, which generate information on a meso-economic level dealing only with a specific branch?

On the topic of comparability of the data, used in existing frameworks, the following questions can be posed:

5. Are the various ways data are collected and treated within the statistical process influencing cross-country comparability (of health care) and if so can this influence be changed in a way positive to comparability of health care data?
6. Are the data on health care cost/expenditure of a selection of European Union (EU) countries as presented in various databases comparable (comparability across countries)?
7. Has the amount of comparability of a set of data elements increased or decreased over time (time comparability) within a country: for example the Netherlands in the period 1972–1999?

The second part of this study presents a number of innovations that can improve the comparability of data on health care. The following questions are relevant in this respect:

8. Recently a number of instruments have been developed that increase the level of comparability, like the Common Comparable Package approach and the

System of Health Accounts. Is there any agreement on the possibilities these instruments offer in improving cross-country comparability?

9. Which opportunities are available to enclose all these new methods in an internationally acceptable framework (like the SNA or ESA)?

The third area, the use of data and the reconciliation of the discussed frameworks aiding comparability, gives rise to the following questions:

10. Which data are used in a national and an international context? Do users realise that data are not always comparable, and if so, does this realisation influence conclusions based on the results? Do they realise that a certain quantity and quality of data is necessary for any research?
11. Is it possible to reconcile the data resulting from the various systems used in cross-country comparability of health care expenditure to reach an acceptable level of comparability? And what should be changed or added to reach a more optimal reconciliation?

The questions listed deal primarily with technical matters on constraints in comparability. Another factor of importance is the willingness on a national level to create comparable data. Of course, several nations must have that will at the same time and execute the work. This (lack of) willingness is not dealt with in this study. This study does not pretend to propound a general theory on comparability. What it intends to do is shed some light on important and less important elements in the comparability of data on health care, and in this way produce a contribution to improved comparability of health care statistics and statistics in general. Furthermore, it aims to make clear that seemingly obvious conclusions resulting from present statistical data are often not in accordance with the reality behind those data.

1.5 *Sequence of chapters*

In the following chapters the questions as mentioned above will be dealt with. In chapter 2 the concepts and notions, used in the international comparability literature, are presented. In the next chapter (3) the SNA as core framework will be treated. In chapters 4, 5 and 6 other existing systems used are presented. These comprise satellite accounts⁷⁾ (chapter 4) as well as the much used (national) health accounts⁸⁾ (chapter 6). Chapter 5 deals with the topic of classifications⁹⁾ by purpose, that could be used in health care comparability analysis. In Chapter 7 the OECD Health Database as an existing framework is discussed. The statistical process is discussed in chapter 8. Chapter 9 contains a comparison of the data on the topic of health care, gathered on the bases of SNA/ESA and Health Accounts. In Chapter 10 the conclusions on the usefulness of these existing methods in the case of comparability are the main topic.

The other chapters in the study are mainly devoted to the recently developed methods to improve the comparability of data. Chapter 11 deals with the method of the Common Comparable Package, as developed by Statistics Netherlands in a project study. In the Common Comparable Package project (CCP-project) a pragmatic solution was searched for (and found) for this comparability problem, based on national sources, national classifications and definitions, to create comparability for a limited number of (six European) countries and a limited number of variables (Mosseveld 1999, page 24). Chapter 12 deals with the newly developed System of Health Accounts (SHA) developed by the OECD in collaboration with international experts and statistical agencies (amongst which Eurostat). In this same chapter the deficiencies of the SHA are presented. In chapter 13 some improvements of the SHA methods are described and applied in the situation of the Netherlands. In a practical example the improved method as well as the original method is shown. Chapter 14 presents some other possibilities for solutions in relation to cross-country comparability of health care expenditure. Chapter 15 deals with the matter of data and data use, both in a national and an international context. In chapter 16 possibilities are discussed to use existing frameworks like National Accounts and national health accounts to reach cross-country comparability. In chapter 17 the conclusions resulting from new developments in comparability are presented as well as the conclusions resulting from the use of data and reconciliation of systems. In the final chapter 18 recommendations to improve comparability are presented.

Notes:

- 1) See e.g. 'Bereidheid tot verandering', Commissie Structuur en Financiering Gezondheidszorg (in short the 'Dekker-Committee', after its chairman prof. Dr. W. Dekker), The Hague, 1987 (in Dutch).
- 2) The League of Nations Health Bureau was preceded by an International Office of Public Hygiene, created in 1907, dealing with public health. The League of Nations overtook this function.
- 3) Health Monitoring Programme (HMP): Council Decision 97/1400 of June 1997 adopting a programme of Community action on health monitoring within the framework for action in the field of public health (1997 to 2001).
In May 2000 the Commission of the EU adopted a proposal for a new Community programme on public health for the period 2003–2008. In 2003, after agreement by Parliament and Council, it is being implemented.
- 4) To prevent misunderstanding it is stated that at no point OECD has indicated these data to be comparable.
- 5) These six factors are: 1. Differences in health care supply; 2. Differences in the definition of health care expenditure; 3. COI does not equal total health care

- expenditure and differs across countries; 4. Differences in data and methods employed in COI studies; 5. Differences in health care utilisation and costs, and 6. Cross-national variations in medical practice.
- 6) The reason is that this study is primarily dealing with economic problems; and outcomes (outcome measurement) are dealing with quality aspects of actions, involving value judgements, received by clients.
 - 7) The main feature of satellite accounts is a seamless link between the aggregates of the satellite account and the figures of the branch as described in National Accounts.
 - 8) The emphasis of (national) health accounts is on the complete description of all activities relevant for the branch.
 - 9) Classifications by themselves are not existing frameworks. Classifications are tools used to order a/o activities, units and transactions and are a prerequisite for comparisons. Functional classifications are important for comparing results both inside and outside SNA/ESA, but are difficult to fit into the general structure of the study. For this reason a separate chapter on functional classifications is included.

2. *Terms and definitions used in the title*

2.0 *Questions*

What does comparability mean? Can any related terms other than comparability be used?

Comparability only has a meaning in a certain defined context; the context of this study is health care. What is health? What is health care? Do any models exist for health and health care? Are these models comparable or completely different from each other?

The last set of questions in this chapter relates to the phenomenon to be examined in the health care area. So what is expenditure? Is it the same as costs?

2.1 *Introduction*

This section elaborates on the terms used in the title of this study: International Comparability of Health Care Expenditure. This means that the term comparability (as well as some related terminology) is described, both in a national and an international setting, as well as the terms health and health care and expenditure (together with related terminology).

I realise that the questions concerning comparability presented in this chapter are formidable and cannot be addressed exhaustively. So by their very nature the answers provided are to a certain extent selective.

2.2 *Comparability and related terms*

The most important term used in this study is comparability. To make any statement possible on comparability, the term has to be clarified. Many articles and books have been published on the subject of comparability in health care statistics. A distinction can be made between more descriptive and empirical articles on the one hand, and more theoretical, econometric articles on the other. The present study uses the term comparability as set out in the empirical and descriptive articles, which emphasise the data and how they can be used to describe similarities and differences between the various systems. In the more econometric articles the data are used in a model, without the actual data being questioned.

A number of terms are used in relation to comparability issues. Often comparability is seen as one of the components of statistical quality (Eurostat 1999a, page 4). According to this report, quality can be defined with reference to the criteria: relevance of concepts, accuracy of estimates, timeliness and punctuality, accessibility and punctuality in dissemination, accessibility and clarity of information, comparability of statistics, and coherence.

The above breakdown of quality is not unique. Other organisations (see Depoutot 1998b, page 5) have proposed other breakdowns for the joint use of statistics. Almost every breakdown mentions comparability over time and place and comparability between data from various distinct sources.

Having worked on quality programmes for several years, Statistics Netherlands uses the encompassing term coherence instead of the more limited term comparability to determine the quality of the statistics and the statistical process. Both the statistical processes and the results (in terms of data) are compared to determine the quality of the products for the user of the statistics as well as for the producer (see Eurostat 2000a).

Eurostat states that comparability can be seen as a special case of the encompassing term coherence. According to Eurostat, "Coherence relates to sets of statistics and takes into account how well the statistics can be used together." (Eurostat 1999a, page 4).

Although comparability may be defined as a special case of coherence, there is no agreement on how the restriction of comparability should be made with regard to coherence. Some restrict the use of the term comparability to statistics relating to the same characteristic, while coherence relates to different characteristics (Depoutot 1998a, page 6; Depoutot 1998b, page 6). Linden (Linden 1999, page 4) is more precise: in comparisons only one of the components may vary while all others remain constant. For others comparability is restricted to different populations while coherence is applicable to statistics for the same population. Eurostat (Eurostat 1999a, page 5) states that "... by comparability we mean that data for different populations can be legitimately put together and interpreted in relation to each other and/or against some common standard."

In the discussions on international comparability the distinction is even stronger (Depoutot and Arondel in Eurostat 1999a, page 5): "International comparability covers discrepancies between the national definitions and the EU norms and discrepancies due to the measurement process. International comparability insists on the links between statistics from different Member States, when coherence and completeness insist on the links between different statistics of the same statistical system."

One of the key elements in this line of reasoning is the term EU norms or more globally norms. This term is not described by Depoutot, but generally speaking norm could mean a kind of standard or reference point, identical for and accepted by everybody. Events or variables to be compared have to be able to be converted from

their original format to the agreed standard. The more successful this conversion is, the more comparable the results will be. The Common Comparable Package (CCP) project studied a few possibilities, one of which was 'conversion, or indirect comparison'. When standards are lacking, references have to be developed. In one of the consultations with the expert group, two versions of a so-called reference hospital or standard hospital were discussed. In the construction of one of the versions of the reference hospital theoretical elements derived from hospital definitions were used, while the other version was based on empirical and more pragmatic elements. In the discussions on this topic it became clear that it was virtually impossible to create such a reference or standard because of differences in guidelines in the various countries, health policy requirements and the lack of enough detailed data (see Mosseveld 1996, page 16). An alternative approach, the CCP approach (see Chapter 11) was adopted and used to create a reference or standard.

Depoutot (Eurostat 1999a, page 5, 6) proposed that the term comparability be used when dealing with a single characteristic, and coherence when different characteristics are involved. According to Eurostat "... facing the statistics and the usual meta-information for several countries ... statistics are comparable if the user can derive from these the message that is embedded in the set of individual data and the proposed angle of observation." (Eurostat 1999a, page 7). Comparability aims at "... measuring the impact of differences in applied statistical concepts and definitions ..." (Eurostat 2000b, page 13). "Systematic deviations in definitions and methods used for the building of statistics ... should be ... provided to the users." (Eurostat 2000b, page 13). In this description of comparability it is obvious that only the producer objective is taken into account. From the viewpoint of the users, comparability is related to concepts and definitions that may deviate from the ones the data producers use. "The user objective in socio-economic analysis is to explain real phenomena [in which analysis] the relevance of the statistical concept is of main interest." (Eurostat 2000c, page 3).

Terms related to comparability include harmonisation and standardisation, two terms used very often in the EU with regard to data. "Harmonisation is used for any process leading to better comparability or better coherence. It includes any effort towards better comparability of statistics relating to a given characteristic or towards a better coherence of statistics relating to different characteristics." (Eurostat 1999a, page 7).

Standardisation is one of the tools for harmonisation, for example the adoption of common definitions, classifications, etc. Standardisation is the elimination or strong reduction of a factor leading to non-comparability or incoherence. Full harmonisation requires complete data standardisation but harmonisation can be obtained by other means as well, for example adjusting figures for differences in statistical units, definitions, classifications, without standardising these units, etc.

Harmonisation can take place *ex ante* (input harmonisation) or *ex post* (output harmonisation). Both terms (input and output harmonisation) were specifically developed by Eurostat to facilitate comparability of data within the EU. At the international level, the best way to obtain fully comparable data – at least in theory – would be to conduct a single statistical operation in several countries as if they constituted a single country. The same result would be obtained by conducting identical statistical operations in each country with the same sampling design, the same operational definitions of statistical units and characteristics, the same measurement process, the same processing, etc. Problems still remain, like different languages, different institutional settings, different non-response rates, etc. Other factors that might influence the level of harmonised outcomes range from objective factors like the reliability of statistics, to subjective categories like different attitudes towards governments.

An alternative to input harmonisation is output harmonisation. Output harmonisation consists in ensuring that the data collected by different countries meet the same definitions, classifications, etc. without attempting to harmonise the data production process itself: each country uses the data collection and data processing techniques it feels are appropriate as long as the end-result is sufficiently comparable.

Increasingly in the EU, attempts are made to harmonise not only the results but also some aspects of the production process (e.g. Labour Force Survey). This process of rebuilding national statistics to international definitions is not always politically accepted. Moreover, it is time consuming and difficult and can only be implemented in a limited number of cases (see Depoutot 1998a).

Output harmonisation is therefore the most common method of harmonisation used by international organisations, not only Eurostat.

The description of comparability provided by Eurostat and others contains important characteristics: the user and producer perspectives, the angle of observation and the data. Different user groups (like national accounts, ministries of health, international organisations) have different goals and viewpoints and therefore need specific data to accommodate these viewpoints. The object of observation is another important characteristic determining the level of comparability. The level of aggregation on which a phenomenon is described partly determines the results in terms of comparability. Lastly, the data supplied provide the message for the user on the level of comparability.

However comparability is not just described by commonalities in the criteria used in the comparison, but also by the description of differences of the characteristics. Commonalities on all possible criteria result in complete comparability. Complete comparability on all criteria cannot be achieved, and indeed is not necessary. What is necessary is at least commonality of (the set of) the most important criteria. Differences in the most important criteria need to be determined and described. For the most important criteria used in the comparisons these differences need to be

quantifiable, so they can be removed or added to the results. Both for the description of the commonalities and the differences of the objects in the comparison, metadata are indispensable.

2.2.1 *Metadata and comparability*

The major studies comprising cross-national evidence include some remarks on the comparability of the data. Most publications on international statistics contain footnotes to tables listing the main causes for non-comparability. Does the determination of comparability of the data, or the lack thereof, perhaps depend on the judgement of the people responsible for the publication? One way to address the issue of international comparability is to require each country to produce a standard quality report describing and assessing the lack of quality based on a large number of possible factors.

One important condition to create more comparability is the availability of metadata. Metadata provide information similar to that of standard quality reports, although some of the items are not relevant for comparability assessment. Metadata are usually seen as data about data, but they also refer to methods used in the collection and manipulation of data. Metadata are key elements in the assessment of international comparability and cross country analysis. "Metadata are not carriers of 'material' information. They are of a 'formal' nature, only containing references to aspects of information: characteristics or typologies." (European Commission, page 123). A formal definition of metadata is presented by the United Nations (UN): "Metadata [are] data and other documentation that describes objects in a formalized way." (UN 2000d, page 20). Statistical metadata are defined as "... metadata describing statistical data" (UN 2000d, page 32). Although this definition is correct (just as the definition "data about data") it lacks context and specifics. Statistical metadata are defined as "... descriptive information or documentation about statistical data, ... [that] facilitates sharing, querying and understanding of statistical data over the lifetime of the data." (UN 2000e, page 1).

A very complete system of metadata, in accordance with ISO/IEC International Standard 11179, is presented by the Australian Institute of Health and Welfare in the National Health Data Dictionary. The metadata system contains items on identifying and definitional attributes, relational and representational attributes and administrative attributes (see AIHW 1998 pages 464–467 for a full description of all the items used). Although the Australian system is very complete, it has been argued that there is more to metadata than the items mentioned in these attributes. Metadata can also refer to the environment from which statistical data are derived, and this aspect is missing in the Australian system.

According to the Organisation for Economic Co-operation and Development, OECD (country descriptions in the Health Database) and the CCP project (country

profiles), metadata should also "... provide the necessary background information for understanding the numerical data included. A 'health care dictionary' and 'standard country descriptions' are suggested as basic categorical elements of a metadata system. The first takes care of proper defining all data elements included, the latter is intended to describe member states' health care system in a standardised way." (Brückner 1997, page 8). An example of a 'data dictionary' is The International Health Data Reference Guide, a publication providing information on "... the availability of selected national vital, hospital, health manpower resources and population-based health survey statistics ... which is not readily available in published form." (DHHS 1999, page iii).

The same kind of addition to metadata is made by Sundgren, when he introduces specific and global knowledge. "Specific knowledge denotes meta-information and meta-information-related functions associated with individual systems, production systems and retrieval systems, that is, meta-information of a relative local nature." (Sundgren 1993, page 18). "General knowledge denotes meta-information and meta-information-related functions of a more global character, for example knowledge about how statistical surveys and information systems are to be designed, operated, and evaluated, which we may call handbook knowledge, since it is often documented in the form of handbooks, manuals and guidelines. Another type of general knowledge is encyclopaedically knowledge, that is, knowledge of the type documented in dictionaries, encyclopaedias, and thesauri. Yet another type of general knowledge concerns standards, contents-oriented as well as dealing with representation formats. Finally there is a subcategory of general knowledge that is concerned with and contained in software products." (Sundgren 1993, page 18). The distinction between specific and general knowledge is reflected in the terms 'local metadata' and 'global metadata' (see e.g. Sundgren 1993, page 41).

Footnotes to tables, sources used, definitions and methods (collecting data, calculating and estimating totals, etc.) are of prime importance for users to interpret the data, especially the differences between countries and deviations from internationally agreed standards (if available). The availability of a set of metadata also provides useful information for statistical agencies, the suppliers of the statistical information. It may encourage countries to start using best practices with respect to the guiding principles for compilation of statistics as well as the use of international standards and classifications. It goes without saying that the compilation and use of metadata are important for all international organisations, especially the standards and classifications (see a/o. OECD 1998a and OECD 1998b). Examples are the SNA (System of National Accounts), NACE (Statistical Classification of Economic Activities in the European Community), ISIC (International Standard Industrial Classification of All Economic Activities), ICD (International Classification of Diseases).

For Depoutot (Depoutot 1998a), on the other hand, collecting metadata is a very time-consuming task, which can never be fulfilled because of constant changes in

national methodologies. He therefore reasons that such a set of metadata is of very limited use to users.

Sundgren's analyses do not support this; he concludes that metadata exert a large influence in increasing comparability of data over time and across space. This view is also shared by those who have been running the OECD Health Database for the last two decades. Sundgren distinguishes two ways of describing comparability: directly, i.e. anticipating which concrete comparisons users would like to make; and indirectly, where the documented characteristics are related to standards. The direct way might be feasible for comparisons in time but very difficult for comparisons across space. The indirect way, on the other hand, depends strongly on the existence of standards (see Sundgren 1993, page 25). This brings us back to one of the prerequisites in the quality description: the existence of standards.

In my opinion collecting metadata is not a once-only event, but a never-ending continuum. The metadata need to be renewed on a regular basis. Moreover without metadata it is virtually impossible to rate the level of comparability of data; metadata are a prerequisite to distinguish differences between data and hence the level of comparability. Metadata were the core business of a European Commission project ¹⁾, funded under the first Health Monitoring Programme, whose results are the basic elements which made it possible to start collecting data on the European health care systems with a clear idea of the level of comparability of these data, and hence the quality of the results.

2.3 *Health and health care*

Everybody knows what health is and uses the term widely. Like many important things in life, health is viewed from a negative angle: the lack of health, and how to restore it. Until the beginning of the twentieth century the government responsibility with regard to health only related to communicable and acute diseases. Health was seen as a bio-medical process, a view that evolved in the late nineteenth century with the growth of scientific medicine and complete public faith in the healing power of medicine. The approach was fostered by advances in bacteriology, immunology, surgery, diagnostics and pharmacology.

In the course of the twentieth century the content of the notion of health and the related notion of health care, and its political implication for government responsibility, was gradually enlarged and changed. The Dutch government only officially recognised its responsibility by law in 1983, by adding article 22.1 to the Constitution, which stated that "... the government will take measures to promote the health of the people" (see e.g. van Son 1989, page 25). After a long period of regulations and laws, government responsibility was finally recognised. The first regulations and laws mostly concerned the promotion of hygienic conditions and, later, safety at work. Health care itself became a topic of policy only in a much later

period, although the medical laws of Thorbecke (June 1865) were directed at health care, i.e. the qualifications and abilities of health care professionals and the supervision of their practices.

Starting from the bio-medical approach, in one direction (content wise) more attention was put on psychological and social issues. Mental illness was a grey area between health and illness, a grey area that was to become more and more important. On the other hand, in this approach only health professionals could heal and recognise disease and decide on the treatment. This last condition started to change in the last decades of the twentieth century, with the development of proto-professionalisation ²⁾. Proto-professionalisation can be described as "... the penetration of the attitude of a professional occupational group in the layman's attitudes" (Geurts 1992, page 4). In the domain of health this means that every layman absorbs knowledge about health and ways to restore his or her health, and puts forward his or her view in discussions with professionals.

All the characteristics concerning the content of health can be found in the definition of health of WHO (World Health Organisation) of 1948. According to WHO ³⁾ "... health is a state of complete physical, mental and social well-being, and not merely the absence of disease or infirmity." This definition of health is widely used, but is not dynamic and difficult to use in practice.

For WHO, health is a state, not a dynamic process. According to other definitions, a dynamic process implies "... being able to fall sick and to recover ..." (Canguilhem 1989, page 199). In a dynamic process health can be seen as a composite "... of the separate concepts of function/dysfunction and prognosis ..." (Fanshell and Bush 1970, page 1033). The concept of function/dysfunction can be defined as a state or a situation ranging from the theoretical state presented in the WHO definitions of health to death (Fanshell and Bush 1970, page 1029, 1030). Health is "... not only the ability to function now but the outlook for future functional ability." (Fanshell 1972, page 322). Whatever starting point is taken for the definition of health will have its influence on the description of the term health care, the boundaries of the system and the consequences in financial terms.

In putting the definition of health into practice in the context of health care, we encounter the same aspects. In the recent World Health Report 2000, WHO defines the health system as "... comprising all organisations, institutions and resources that are devoted to producing health actions. A health action is defined as any effort, whether in the personal health care, public health services or through intersectoral initiatives, which primary purpose is to improve health." (WHO 2000, Overview, page XI).

Looking at the definition or description of the health system as supplied by WHO ⁴⁾, in my opinion what WHO describes corresponds more with the health care system than with a health system, as they restrict the definition to the collection of providers of health care; the autonomous or exogenous factors, often called determinants, are

not part of the definition. These autonomous factors, however, are crucial elements of a health system.

2.3.1 *Health statistics*

Neither the health of an individual, nor that of a population can be seen in isolation. Health is influenced by the social system, technology and environment, as well as by lifestyle and hereditary factors. All experts in the field see these determinants as influencing health and playing an important role in the health system. The various models of health and health care all contain these elements. Three Dutch examples of these health models are presented below, all based on Lalonde's ⁵⁾ model: one published by the Raad voor de Volksgezondheid en Zorg (RVZ 2000), one developed by Statistics Netherlands (1982 and 1999) and one developed by the National Institute on Public Health and the Environment (RIVM). A slightly modified version is published in the Public Health Status and Forecasts (1997).

Lalonde (Lalonde 1974) developed a new concept, envisaging "... that the health field can be broken up into four broad elements ..." (Lalonde 1974, page 31), the four elements being human biology, environment, lifestyle and health care organisation. The human biology element includes all aspects of health, both physical and mental, which are developed within the human body as a consequence of the basic biology of man and the organic make-up of the individual. The environment category includes all matters related to health which are external to the human body and over which the individual has little or no control. The lifestyle category in Lalonde's concept consists of the aggregate of decisions by individuals which affect their health and over which they more or less have control. The fourth category in the concept is the health care organisation, which consists of the quantity, quality, arrangement, nature and relationships of people and resources.

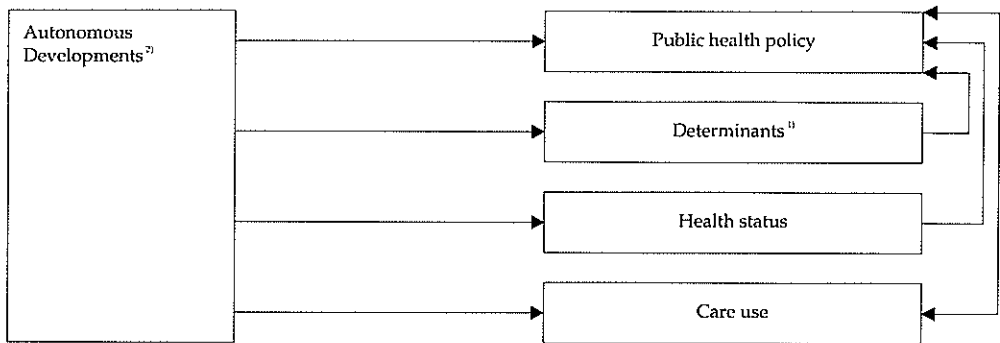
So all four elements, not just the health care organisation, were presented as factors determining the level of health. Any health problem could be traced back to a combination of these four elements. In Lalonde's model, the health status of the population is the focal point and the relations between the determinants and health are visualised.

Lalonde's concept became the basic model for other organisations. An adapted version of this basic health model (see RVZ 2000, page 6) shows clearly the direct influence of the four determinants on public health, and the influence of health status and health care on each other.

The second example given here is the health model developed by the National Institute on Public Health and the Environment (RIVM), first published in the Public Health Status and Forecasts of 1993. In the study of 1997, health care use is added to this model. Figure 2.1 presents the model in its simplest form.

The model shows that the health status is fixed by determinants and that this status influences the use of the care system and its expenditure. Health status,

2.1 Model used in Public Health Status and Forecasts



¹⁾ Determinants: Health Care (physical and mental), Prevention.
(RIVM 1997, page 84) Exogeneous (physical environment, social environment, life style).
Endogeneous (heredical, obtained).

²⁾ Autonomous Development: Demographic factors/developments.
(RIVM 1997, page 134) Macroeconomic developments.
Social cultural developments.
Medical technological developments.

Source: RIVM 1997.

determinants and use of the care system influence public health policy, which in turn tries to influence the health status through its determinants. This process is influenced by autonomous or exogenous factors (like demography, macro economic factors, social cultural factors and medical technology).

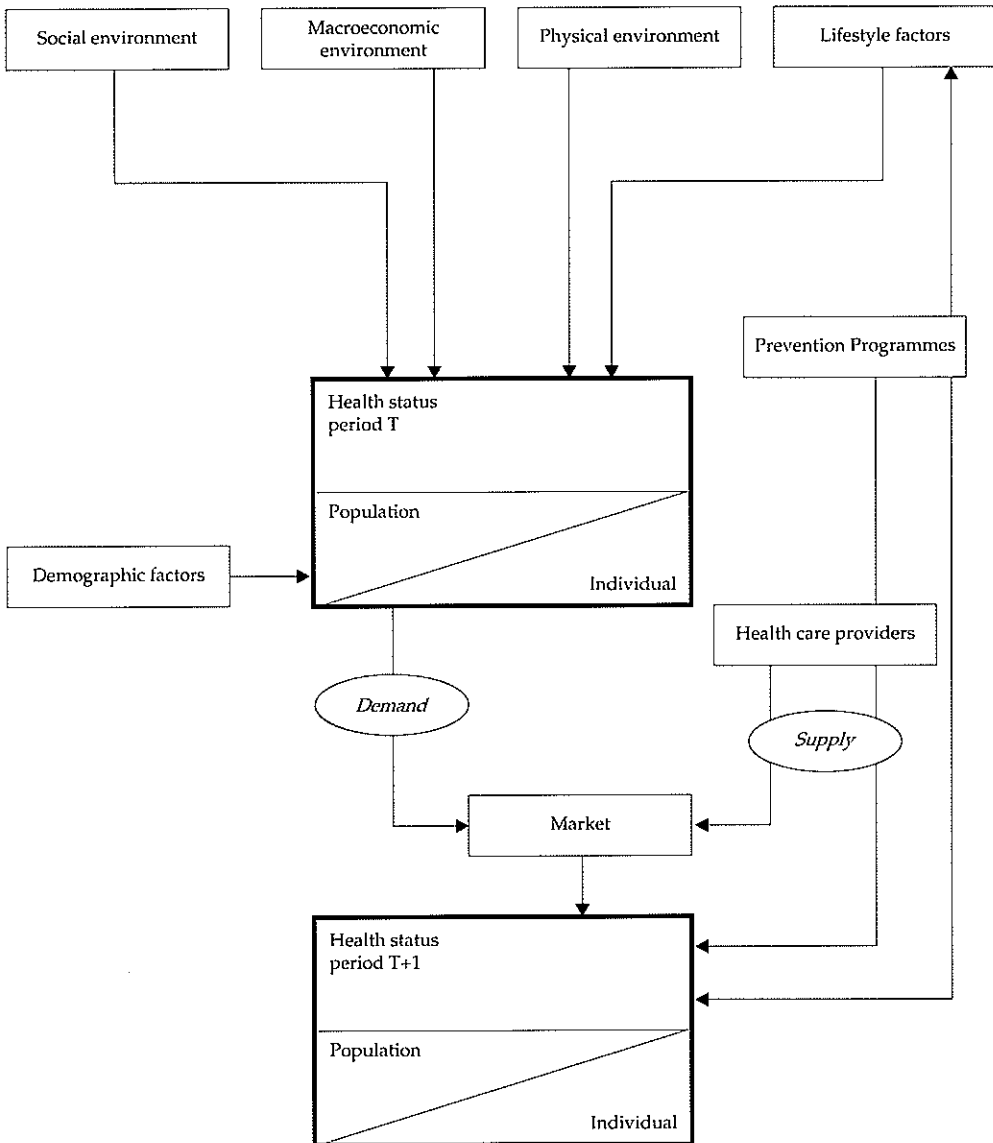
Lalonde's model was the starting point for the development of a health model by Statistics Netherlands (for a first version see Bonte 1982). In this health model, the term health status is used instead of health, as the former is much easier to put into practice: a status can be measured with the aid of variables. For the other variables influencing the health status the RIVM approach is used, as the terminology used in their Public Health Status studies are logical and widely accepted in the Dutch community. By using these terms the statistical office uses the same language as the experts outside the office.

Because the health care system is part of this integrated model of health, the same relations between these variables are valid for the health care system. The health and health care units of Statistics Netherlands aim to describe and present information on all the above-mentioned topics. The presentation of coherent statistical information on the health status, the services rendered by the health care providers, the way the health care system operates and the relations between health status and health care system are a primary target. The model (including welfare statistics) is presented in a report by Statistics Netherlands (CBS 2000a, chapter 2).

The figure as published in CBS 2000a was built on two separate parts: one containing health and health care, the other containing welfare or well-being. The figure

presented in that publication was by necessity a compromise between health (care) and welfare/well-being. Figure 2.2 presents only the health section of the combined figure as published.

2.2 Overview of relations regarding health



Source: Derived from CBS 2000a.

All elements mentioned in this section are part of the model used by Statistics Netherlands to describe the system of health care. Because the model is used for an economic description of the health care branch, special attention is paid to the role of the demand and supply aspects and the market as the meeting place for supply and demand. Central to these economic relations are production and output on the producers' side, and consumption expenditure on health care on the consumers' side.

Although it might seem that all three models are completely different, there is just as little truth in this statement as in the statement that all three are exactly the same; they are in fact all based on the same notions using the same parameters. The differences lie in the value attached to the various components, and in which relations receive specific attention.

2.3.2 *Health care statistics*

Three areas can be distinguished in statistics on health care (see Figure 2.2):

- statistics on health and sickness (and its determinants) especially at the population level, represented by the health status box;
- statistics on the volume of services requested and supplied (meeting on the market);
- statistics on the total health care system in an economic sense, i.e. statistics on providers and prevention programmes, e.g. on material, money and manpower.

Basically, it can be stated that health care includes all activities concerned with health as defined above, either in the classic definition of WHO (1948) or the more dynamic ones (e.g. by Canguilhem and Fanshell and Bush). Linking health and health care to more statistical and economic dimensions puts providers of health care more in the picture. In short, it can be stated that health care providers are all institutions and independent professionals that supply services to people or groups of people with the aim of improving health, or at least to prevent further deterioration of the state of health.

From a statistical point of view, a system of health care statistics can start from the perspective of the care provider, the financing unit or the consumer. In the end, the information collected should lead to a coherent set of data. Most statistical systems start out from either the care provider or the financing institutions, but all systems ultimately aim to provide data on the consumption of goods and services in the health care branch by the population.

The OECD (OECD 2000, page 42) defines the concept of health care as "... the sum of activities performed either by institutions or individuals pursuing, through the application of medical, paramedical and nursing knowledge and technology, the goals of:

- promoting health and preventing diseases;
- curing illness and reducing premature mortality;
- caring for persons affected by chronic illness who require nursing care;
- caring for persons with health-related impairment, disability, and handicaps who require nursing care;
- assisting patients to die with dignity;
- providing and administering public health;
- providing and administering health programmes, health insurance and other funding arrangements."

The bio-medical approach to health is included in this definition, but the more dynamic parts of modern health are not directly visible (see Canguilhem, Fanshell and Bush). The same is true of the definitions of health care and health care institution used by Statistics Netherlands for the description of the area of health care in both the institutional approach (SNA) and in a more functional approach.

In the functional approach the area of health care is described as follows (CBS 1994, page 9): "Health care concerns the supply of goods and services in the area of medical, paramedical and nursing care.

These goods and services:

- are provided for human beings suffering from diseases, disabilities and/or limitations of a physical and/or mental nature;
- are related to prevention, diagnostics, treatment and nursing/caring;
- are produced by trained experts and/or companies (or parts of companies) set up for this purpose."

For the institutional approach in the national accounts a definition derived from this functional definition is used. The following criterion for the business units is added to the functional definition mentioned above: 'These units should operate as independent economic units and, classified by main activity, should belong to this health care class in the classification.'

Although not explicitly part of the functional description, expenditure by the government for policy, administration and management as well as expenditure by private and public insurance companies in this area are included in the functional approach.

Of course, neither policy, administration and management, nor insurance companies are part of the health care branch in an institutional setting.

2.3.3 *Production boundary in relation to health care*

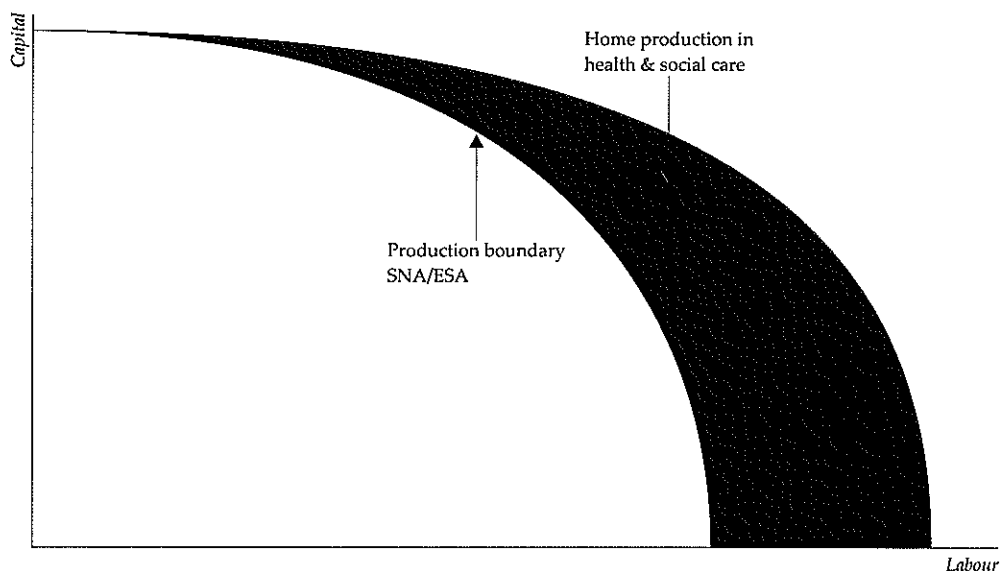
All the descriptions above do not give a solution for the production boundary problem.

Abel-Smith already mentioned five different border-line problems with regard to health care in 1963. The first is the distinction between environment and public

health, the second between normal body care and health care, the third between ordinary child welfare and promotional health care for children, the fourth relates to institutional care and sick persons admitted in these institutions and the fifth is the distinction between education and research on the one hand, and health care on the other (Abel-Smith 1963, page 26). In 2003 these border-line problems or boundary problems are still more or less valid, although an acceptable consensus has been reached for the third and fourth problems (child health promotion and social institutional care).

One of the important problems, still existing in various statistical systems in relation to health and social care concerns the production of health and social care by private households in their role as production units (see Figure 2.3).

2.3 Production boundary



In theory three possible solutions are feasible: including all production from households, including only the production of households for which financial flows can be identified, or including no household production in this area at all.

Again in theory, all the production of health care at home could be included in the production boundary, whether this production is for own final use or for the final use of other households, paid or unpaid.

In a national accounts approach all production of goods (including home production) is included in the production boundary, whether produced and

delivered to other institutional units or retained for own final use. Home production of services within the production boundary is limited to imputed rent for owner-occupied dwellings and personal (household and butler) services. Home production of health care services for own final use is not included, not even if financial payment is involved. The national accounts principles present a clear boundary but leave out a more or less substantial part of production of health and social care services produced at home with or without any financial compensation. The size of the 'disturbance' depends on the level of 'unpaid' assistance in the various countries.

Classifications and resulting comparability also give rise to problems. In various countries the boundaries of the systems are set along different lines. In national accounts, for example, the provision of pharmaceuticals and therapeutic appliances is not part of the health care system, but part of retail trade. In almost all (more functionally oriented) health accounts systems, the supply of medical goods is seen as a part of health care.

Some systems include matters like education and training, research and development, food hygiene and drinking water control, environmental health, services in kind to assist patients, and health-related benefits, while in other these are excluded. Other areas in the health care boundary include all kind of social activities like caring for the mentally handicapped, caring for the elderly, etc. All the areas listed above can be grouped together under the heading providers of services and country specific organisational differences.

A related issue, but not to be grouped under the boundary heading, is the treatment of capital formation in the health care branch. Depending on the financing systems of the branch, either capital formation is totally or partially included or excluded and depreciation (as a proxy for using the capital) and interest payments are included or excluded in the statistical data.

Even without looking across the border of a national system of statistics, comparability is not guaranteed, as it may be prevented by alternative definitions and descriptions.

2.4 *Cost and expenditure*

The last two terms to be described are cost and expenditure. In business economics the distinction between these two terms is important; they are not identical but may be two sides of the same coin. According to Van der Schroeffer (1970, page 7), cost is the necessary economic value to be sacrificed in the production process, while expenditure is the financial equivalent to be sacrificed to gain inputs (i.e. all

production factors: labour, machinery and equipment as well as inventories) for the production process. Expenditure and cost of the same item may relate to different periods (e.g. expenditure on raw materials and the use of these same materials). Expenditure is by definition not identical to cost, the amount of expenditure in a certain period may be larger or smaller than the costs of the production process. Investment outlays are part of expenditure but not of the costs of that period; depreciation, on the other hand, is part of the cost of the production process but does not constitute an expenditure.

In this study expenditure is set down as the financial representation of society for having and maintaining a health care system and not the cost, in a business economic sense. The cost for the society of having a health care system arises from the fact that goods and manpower and other production factors are used to produce health services and the associated operation and management. The process of measuring the monetary costs of producing health services consists of valuing the resources needed to provide them.

In accounting, especially national accounting, it is useful to distinguish between the running costs, i.e. costs directly associated with the present production of services rendered, and capital costs, i.e. costs incurred for the production in periods to come. Capital costs, however, is not a correct term. What is meant is expenditure for capital goods, and expenditure is not costs. The costs of capital goods to be included in the production cycle is the equivalent for the use of the capital goods, the depreciation.

By and large two possible systems can be distinguished: one in which the government supplies investment goods directly free of charge, or supplies the financial means for investment, which means that suppliers of services do not include depreciation in their accounts; and the second whereby all health care producers have to finance their investment out of their receipts, in which depreciation is included in the financing of their cost structure.

There are two ways to determine the complete expenditure of society or the cost for society for maintaining a health care system. The first one is by adding up all expenditure by the financing units (government, insurance companies, businesses and consumers) including investment subsidies and investments in kind. The second way is to add up all costs incurred by all the producing units including depreciation but excluding payments for investments.

Both methods result in a total amount for having a health care system, one in the total expenditure, and one the total amount of costs.

The problem is not the choice to be made or the distinction between the terminology used, but a more pragmatic one: how and where in the economic system can these expenditures and costs (best) be measured? In theory, there are various options for the point of measurement. For the health care system in a narrow sense (i.e. the supply or delivery of care to the users) the ideal point of measurement is the

consumer sector (private as well as government) as the ultimate beneficiary and the ultimate payer. However, it is virtually impossible to measure all the expenses of the consumption sector for health care. The data that can be measured at the consumer end (mostly expenditure on premiums) is not the information needed (payments on an accrual basis for the goods and services consumed).

The second best option provides a very good alternative: the care providers. Being business units, the providers are almost always in the possession of a reliable administrative system, from which enough and consistent data can be derived. Data on the expenditures as well as the costs (more or less) and detailed information on the receipts are almost always available.

Data derived from the provider administrations alone do not cover the total health care system, though. Additional information is needed on the expenditure or cost incurred by organisations in the area of control, policy and management. This information can be derived from the administrative systems of these bodies.

A related problem concerning expenditure is the distinction between accrual (transaction value) base and cash value base. What is needed is the representation of the expenditure related to transactions during a certain period and not the payments for goods and services in a certain period. Payments in a period may relate to that same period, but also to the previous or subsequent period. For statistical purposes, accrual based data are usually requested.

From a national accounting point of view the production value of the health care branch should be measured. The counterpart of the production is consumption, in terms of consumption expenditure (both individual and collective).

2.5 *Final remarks*

Returning to the questions posed at the beginning of this chapter it can be concluded that comparability is a very difficult term to define unambiguously. According to Eurostat and others, statistics are comparable if the user can derive from them the message that is embedded in the set of individual data and the proposed angle of observation. Different user groups and producer entities have different perspectives, different angles of observation and thus need different data. Comparability for one type of user employing one type of observation angle does not imply comparability for the same user employing another angle or for other users.

Health and health care, the subjects of comparability, are apparently easier to define. According to WHO, health is a state of complete physical, mental and social well-being, and not merely the absence of disease or infirmity. Supplementing this definition with the notions of being able to fall ill and to recover adds a time element needed (at least for statistical purposes).

OECD defines health care in a country as comprising the sum of activities performed either by institutions or individuals pursuing the goals of health (restoring health, curing diseases, caring for the sick, etc.) through the application of medical, paramedical and nursing knowledge and technology. Statistics Netherlands describes health care in an economic sense as concerning the supply of goods and services in the area of medical, paramedical and nursing care. The primary targets of a coherent statistical information system on health care should be information on the health status, services rendered by health care providers, how the health care system operates and relations between health status and health care system.

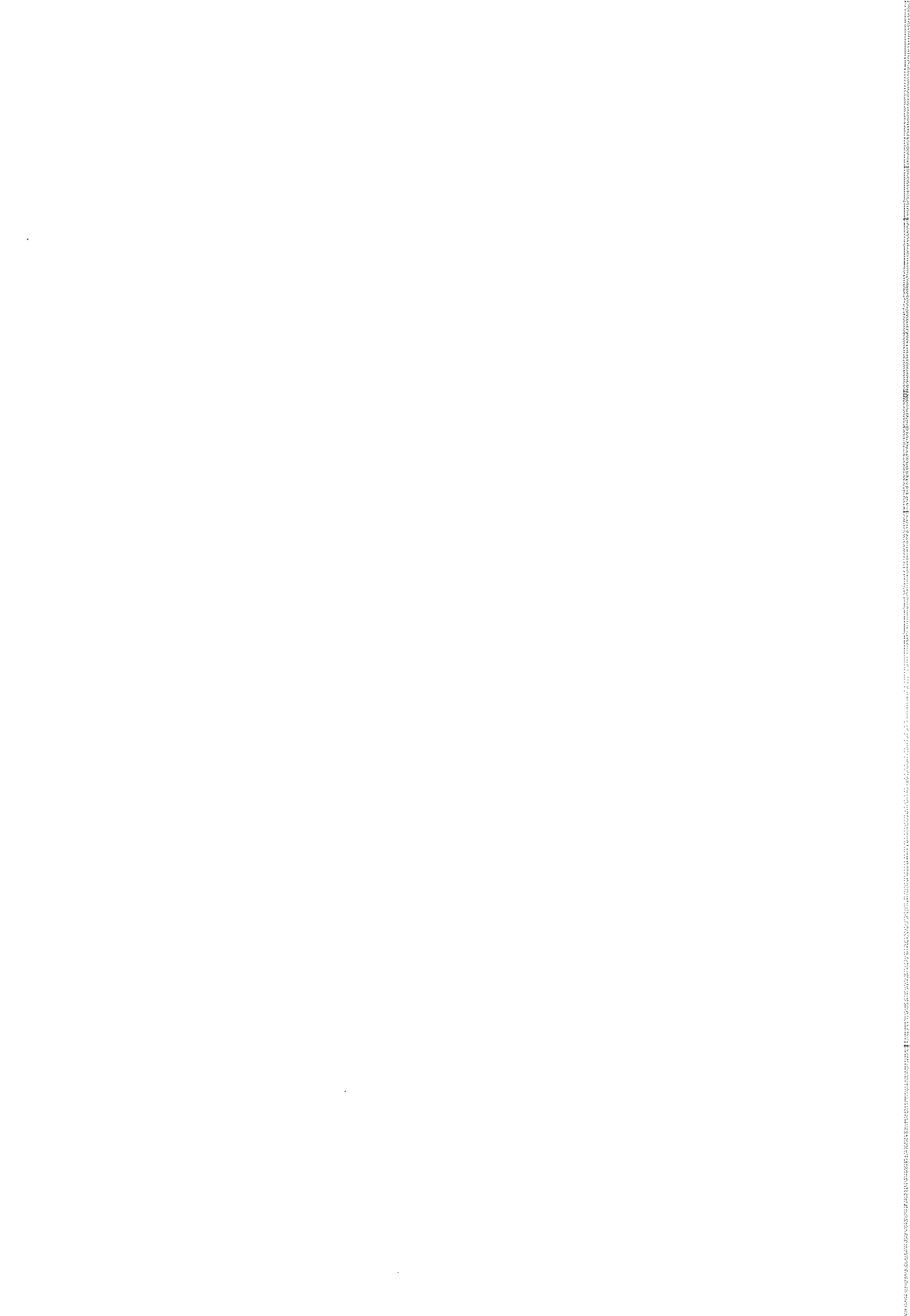
Unfortunately the boundary of the area of health care is not identical for all countries, and this adds to the problem of measuring how much is spent on health care. Cost is the necessary economic value to be sacrificed in the production process. Expenditure is the financial equivalent to be sacrificed to gain inputs to be used in the production process. Health care expenditure is determined as being the financial representation of society for having and maintaining a health care system.

Having explained the terms in the title, and presented comparability as a problem in relation to the boundaries of economic production, it is time to turn to statistics. Comparing data is only relevant in an agreed setting. In economics one of the best settings is the framework of national accounts. The System of National Accounts will be discussed in Chapter 3.

Notes:

- 1) EUCOMP, Towards Comparable Health Care Data in the European Union, on behalf of the European Commission prepared by the North Eastern Health Board, Ireland, Voorburg/Heerlen, 2000.
- 2) Proto-professionalisation was introduced as a special topic in the Dutch Health Interview Survey in 1991. Results on this topic are presented by Statistics Netherlands in Maandbericht Gezondheidsstatistiek, Dec. 1992).
- 3) Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19–22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2 p. 100) and coming into force on 7 April 1948. This definition has not been amended since 1948 (Source: <http://www.who.int/about.definition/en>).
- 4) On page 5 of the same report a somewhat deviating description is provided. Here it is stated that “This report defines a health system to include all the activities whose primary purpose is to promote, restore or maintain health.”
- 5) There have been other models, but since 1974 the Lalonde model has had the most influence on policy analysis and statistical developments in Europe.

*Part 1: Description and analysis of
existing economic frameworks*



3. System of National Accounts/European System of Accounts (SNA/ESA): relations of aggregate economic activity

3.0 Questions

Which criteria and starting points of the System of National Accounts (SNA) and the related European System of Accounts (ESA) are relevant for comparability?

What do these criteria and starting points, in practice, mean in relation to comparability (at a lower than aggregate level)?

Do (recent) developments provide any opportunities to advance comparability, especially of sub-aggregates?

3.1 Introduction

Many of the articles discussed in chapter 2 show clearly that comparability of data is an important issue in the European Union. It is further obvious that they do not yet contain a theory on comparability, only basic ideas for theoretic analysis. In the absence of a theoretical framework, the only possibility is to look at frameworks, especially internationally accepted frameworks and their relation to comparability.

The System of National Accounts or SNA (and the European System of Accounts or ESA) is a basic system used in the modern statistical foundation of the economic theory. This chapter discusses these two (related but not identical) systems, as well as their evolution and how they may provide starting points and solutions for the problems of comparability.

3.2 SNA/ESA: background

The origins of the construction of the SNA/ESA are based in economic theory. To understand this, we need to know how the various economic entities interact. It was J.M. Keynes who presented this insight in his *General Theory* (1936), making it

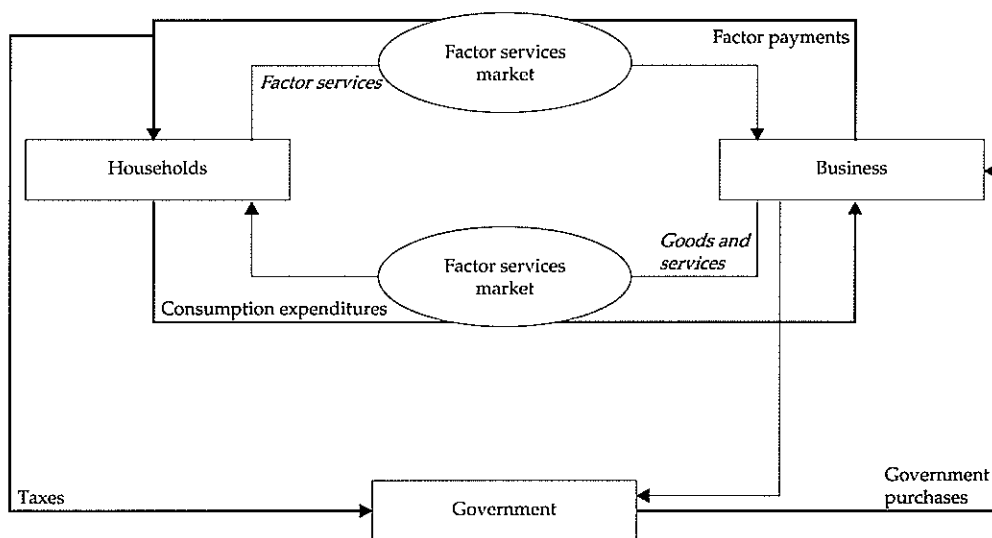
clear how essential it is to have a method to obtain an overall picture of the level and rate of change of the nation's economic activity. Society wants to be able to tell what the resources of the economy are producing and how the output is used. This is the function of the national income and product accounts. The national accounts express the level of performance of the economy in an aggregate sense and allow for comparisons with past levels.

Economic theory requires data to support and substantiate it. Data are also needed to validate the outcomes of government programmes. Another important point in the development of the SNA is the fact that the problems the world faced in the 1920's and 1930's were of a global nature. Global problems demand global solutions and uniform systems of data handling.

The main impetus for the development of a systematic approach to national accounting was the practical demand for information about the working of the economic system as a whole and the way in which its various parts interact. "The practical need arose largely from the great depression of the 1930's and the subsequent problems of economic mobilisation and war finance in the second world war. In the post-war period the information was desired to throw light on problems of economic reconstruction and development, and more generally, for assessing economic change as a background for economic decision-making in connection with public policy." (SNA 1953, page 1).

Data in themselves are not enough; a theoretical superstructure is necessary within which the picture presented by the accounts can be analysed. The usefulness of

3.1 Circular flows including government



Source: Derived from Demburg & McDougall 1972.

national accounts lies in the fact that they give a systematic presentation of the major economic flows in the framework of a comprehensive system, and that they facilitate the understanding of the statistical relationships between these flows. The mass of numbers that compose a complete set of national accounts would only be of limited use in terms of measurement if no theory were available by which the record of the past could be used as a basis for predicting the future.

F. Quesnay (1694–1774), the founding father of the Physiocrats, was the first person to present the idea of the processes of money and goods as a cycle. The growth of macro-economics in particular showed clearly that the cycle theory was an excellent way to present economic events. Basically an economic cycle model distinguishes a few sectors¹⁾: companies and consumers, usually supplemented by government and a foreign sector (sometimes savings and investments are added as separate items). Figure 3.1 illustrates the basics of the cycle process (adapted from Dernburg and McDougall 1972, page 27).

The SNA (and the ESA) is built on the notions of the economic cycle theory. The majority of statistical systems in turn are built on either of these two systems. Although statistics on gross domestic product (GDP) are extremely useful, the necessary process of aggregation involves simplification. Thus the single figures representing GDP conceals a large amount of useful information. This loss of information might be seen as the cost of aggregation. Any accounting framework necessarily risks loss of information as the price for presenting a simple and comprehensive picture, whether they are the books of an individual business, or the accounts for the economy as a whole. The problem is, then, to aggregate data to a limited number of meaningful subtotals.

Although the data are rough, they can effectively dispel some of the odd notions about economy people are likely to have. Is government taking a larger bite out of GDP? Which industries are most important? Correct answers to such questions of great significance to businessmen, employers, trade unions, policy-makers, voters and many others.

3.3 *A System of National Accounts*

National accounts are not just a case of adding up numbers. In many individual cases decisions have to be taken on which transactions²⁾ should be included or excluded. Such considerations mean that there are different views on what the contents of GDP should be, even when adhering to the rules laid down by the UN (United Nations). Although most countries follow the UN rules and guidelines in the compilation of their own national accounts, this does not necessarily imply comparability across countries.

The guidelines³⁾ concerning the compilation of national accounts were published by the United Nations as A System of National Accounts and Supporting tables. Up to now the UN has published three major versions of this System of National Accounts (SNA). The first dates from 1953, the last from 1993. Since the first publication in 1953, only two major revisions have taken place: in 1968 and 1993. The 1968 revision was published under the official title A system of National Accounts. The two revisions in 1960 and 1964 were very minor. These new releases not only contained more precise definitions of terms, rules and guidelines used, but also new subjects that had been important in the periods preceding the new releases.

3.3.1 *The SNA: criteria and starting points*

The purpose of a system of national accounting is "... to provide a framework for reporting national income and product statistics which is of general applicability ... with particular emphasis on the development of international standards for concepts and classifications." (SNA 1953, Preface, page VII).

These two very important starting points, criteria, are mentioned in the preface to the first edition of the SNA in 1953 and have been in the centre of attention ever since. The first chapter mentions the need for a systematic description and a coherent picture of the economic structure (SNA 1953, page 1), as well as the construction of a comprehensive system (page 2) although the latter was not yet thought feasible in 1953. Also in this edition: "... need for international comparability should be kept in mind." The international reporting of comparable national accounting data is mentioned explicitly in the 1968 version of the SNA (SNA 1968, Preface page III). A third starting point explicitly mentioned in the 1953 SNA concerns compatibility. Compatibility, although implicitly very important, can only be established by the references made to other classifications, like the International Standard Industrial Classification of All Economic Activities (ISIC) on page 40, and the Standard International Trade Classification (SITC) on page 41, used as a proxy for private consumption expenditure.

Briefly, the basic characteristics of the system of reporting of information mentioned in the report are:

- comprehensiveness;
- consistency;
- international comparability;
- compatibility.

The basic characteristics can be interpreted as the criteria to which a system should adhere. So already in the first edition of the SNA, wisely called 'An SNA' by the United Nations Statistical Office, the criteria for a system were mentioned.

For the problems discussed in the present study, two notions can be distilled from these criteria. The first that comparability is mentioned as an important issue; indeed as a cornerstone of the system. The second that in the actual chapters, this

importance is not reflected in terms of numbers of words and rules devoted to comparability.

It is important to realise that in the SNA comparability is limited to the results at the macro-level (e.g. GDP, production, final consumption). Data needed for comparability at a lower level of aggregation than the aggregate level are less important. Comparability (across countries) of data on production, consumption and other aggregates, collected at the branch level (e.g. retail trade, agriculture, health and social care) call for additional requirements for the branch system (in view of cross-country comparability) and for the connections with the national system of accounts (in view of completeness and consistency).

3.3.2 *The development of the SNA*

Public policy and the use of national accounts in decision-making is not limited to the global economic goals of full-employment, stable price levels, stable exchange rates, etc. Already in the 1953 edition of the SNA possibilities were mentioned for sectoral approaches to economic policy and decision-making. On page 2 the authors state that "... useful conclusions can be derived from such factual information [on e.g. the production amounts in various industrial sectors] without any elaborate analysis. Even a rough quantitative picture ... is valuable for many administrative purposes and for this reason national accounting should be undertaken even ... if they cannot be expected to attain a high degree of quantitative precision." (SNA 1953, page 2). The last part of this sentence – the lack of quantitative precision – would be a very limiting factor for the usefulness of these data for international comparisons, if this was still valid today.

Of course since 1953 the SNA has grown to become a more complete system. The 1968 version introduced extensions such as the disaggregation of the production accounts into input and output accounts for separate industries and commodities, the breakdown of net lending or borrowing into the transactions in financial assets and liabilities of sectors, and the division of income, outlay and capital accounts into corresponding accounts for these categories (sectors in the economy), and the balance sheet accounts for these categories. It also introduced additional classifications for activities of government and private non-profit bodies and transfers of income to furnish more adequate tabulations. And lastly the 1968 version integrated constant price data on the supply and disposition of goods and services into the structure of the system (see SNA 1968, Preface page III).

The 1993 revision also included many new features, most importantly concerning updating, clarifying, simplifying and harmonising aspects. The system needed to be updated because of the evolution of the economy: new services and industries evolved, new markets were created, etc., and because political systems and influences changed. Clarification and simplification were needed because the system needed to remain applicable in economies that were becoming increasingly complex and changing all the time. Lastly the 1993 version incorporates the results of the harmonisation of the system and its components with other international organisations. The result is most obvious in

the balance of payments, as the revision of the SNA was coincided with a revision of the IMF's (International Monetary Fund) balance of payments manual. The SNA and ISIC (revision 3) were brought together so that definitions of kind of activities and statistical units are the same. And ILO (International Labour Organisation) and SNA employed the same production boundary, enabling better comparisons (see SNA 1993, Preface, page XXXIII and XXXIV).

On an aggregate level international comparability of the national accounts data has gained from the improvements in the system – i.e. availability, consistency of classifications, accounts and rules for calculations and estimations – since the first edition in 1953. On a lower than aggregate level cross-country comparability also improved, but the definitions, classifications and rules and regulations governing national accounts are still the limiting factors for the usefulness of these data.

One characteristic or criterion not mentioned in any of the earlier versions but explicitly in the centre of the 1993 revision is the term flexibility, to create "... the means of facilitating international comparisons and encouraging the use of SNA in economies that differ widely ..." (SNA 1993, Preface page XXXIV). One aspect of this is that the classification system and accounting framework may be used at different levels of detail; another is the use of satellite accounts permitting additional concepts, classifications etc.

On the one hand of course a certain amount of flexibility will create greater possibilities for international comparability, as long as at specific levels of aggregation the degree of flexibility does not differ between the various reporting systems. Greater flexibility promotes the use of the SNA and its results, thus reinforcing the possibilities of the SNA to play a central role in statistics.

On the other hand, however, flexibility may prevent comparability of the results. Flexibility applied in practical situations might result in better linking of branches (e.g. health care, tourism) in the central framework, but at the same time generate less comparable results of the branch concerned. So the total aggregate might be more comparable, while at the same time the detailed results are less comparable across countries.

3.3.3 *Functional classifications and functional approach*

The preface to the 1968 SNA mentions the inclusion of new, functional, classifications. It is important to state that functional classifications are not identical to functional approaches. At Statistics Netherlands the functional approach is used to obtain all the information on a given field of interest: tourism, environment, transport or health, for example. All activities in the economy relating to this field are important; statisticians aim for a financial presentation of a complete system of production, consumption and financing, while at the same time attempting to relate non-financial data to the system as well. Functional approaches to health care will be discussed in chapter 6.

Looking at the description of a functional approach, it comes close to what the SNA 1993 calls a satellite account. Such satellite accounts, although in themselves a useful addition to the system of national accounting, do not solve the problem of comparability across nations or regions (see chapter 4).

SNA 1968 introduced the functional classification of inputs. Page 15 reads "... inputs are classified by the component of value added but no attempt is made to classify these inputs further. They may however perform a number of more or less distinct functions in addition to providing a basis for productive activity. Some may be used to maintain recreational and medical facilities, which, in part at least, are of direct benefit to employees though they may also contribute to productiveness." It is clear that function here has the limited sense of supplying additional value used in an intermediate process. In SNA 1968 this is stated as "... to assign the inputs already distinguished to different functional categories, which would involve a further subdivision of these inputs." (page 15). In a second stage decisions are needed on treatment of the expenditures. Three obvious possibilities are available:

- a separate category of expenditures on inputs;
- inclusion in supplements to wages and salaries, thus becoming part of private consumption expenditures;
- a new category of industrial final consumption.

In SNA 1968 no decision is taken on which path to follow; it states that this area requires further study. However, in Chapter 5 of SNA 1968 classifications of the purpose of transactions are presented in various tables (e.g. table 5.3 on page 87–89 classifying the purposes of government; table 5.4 on page 89, classifying the purposes of non-profit institutions serving households). In chapter 6 the classification of households goods and services is presented in table 6.1 on page 105–108.

SNA 1993 devotes a chapter (SNA 1993, chapter XVIII, page 415–420) to functional classifications, functional in the sense of identifying the purposes or objectives of transactors and their transactions. The four functional classifications (individual consumption by purpose, functions of government, purposes of non profit institutions serving households and outlays of producers by purpose) serve three purposes:

- to distinguish between collective services and individual consumption;
- to provide statistics relevant for a wide variety of analytical uses;
- to provide users with the means to recast key aggregates for particular kinds of analyses (e.g. satellite analysis).

In the SNA functional classifications are used to create separations in main macro-economic aggregates, e.g. by sector. Functional classifications are not used to create an accounting system for such a specific area, like health, tourism or transport.

It goes without saying that the inclusion of the functional classifications in the SNA (1968) and the introduction of satellite accounts (1993) are very important. If international agreement in the use can be realised for these two topics, they can both be key factors for increasing comparability. Both systems will be dealt with in other chapters. Neither the availability of these systems nor the existence of functional approaches will, however, solve the problem of incomparability of the results.

3.4 *The European System of Accounts*

In parallel with the evolution of the SNA, and linked with this system, the European System of Accounts, or ESA, was developed. The ESA 1995 (full title: European System of National and Regional Accounts) replaced the ESA 1970 (European System of Integrated Economic Accounts). A slightly modified second version was published in 1979, the main new features being the change to net recording of VAT (Value Added Tax) and the introduction of a ninth chapter on the measurement of changes in prices and volumes (ESA 1979, page 4). It is not that remarkable that the first European co-ordinated economic accounts were created almost twenty years after the first SNA in 1953, and indeed more than 25 years after the first publication of guidelines by the OEEC, Organisation for European Economic Co-operation). It was only after the first SNA was published that the European unification movement ⁴⁾ started and the systematic description of the economies in Europe became important.

3.4.1 *The ESA: criteria and starting points*

Like SNA 1993, ESA 1995 is "...an internationally compatible accounting framework for a systematic and detailed description of the total economy, its components and its relations with other total economies." (ESA 1995, page 1).

ESA 1995 is consistent with SNA 1993 with regard to definitions, accounting rules and classifications. Differences between the two publications are primarily located in the way the system is presented. ESA is more in line with the use within the European Union, which requires greater accuracy in the definitions and the accounting rules. To ensure that the rules are strictly enforced, the Council of the European Union decided to give ESA 1995 a legal basis in 1996 (Council Regulation 25 June 1996).

In addition to the normal uses made of ESA (or SNA), which include analysis of the total economy, development of a total economy over time, and the relationship between economies, for the European Union and its member states, the figures from this framework play major roles in formulating and monitoring social and economic policies. More important, however, is the use of ESA as a guiding principle in the European monetary policy (convergence criteria), the division of funds over the regions of the member states, and lastly the determination of the resources of the European Union itself and its bodies.

Like SNA 1993, ESA 1995 also presents its characteristics in the description of its concepts:

- comprehensive;
- consistent;
- internationally compatible;
- flexible and multi-purpose;
- harmonised with other (social and economic) statistics;
- practical in use, well-established and fixed for a long period;
- different from administrative concepts and focused on monetary and readily observable terms.

Comparability is an important characteristic that might be expected in the ESA, as so many decisions are based on its results. However, comparability is only referred to in relation to compatibility. "International compatibility of concepts is crucial when comparing statistics for different countries." (ESA 1995, page 2). Both in the 1970 edition (Foreword, page 1) and the 1995 edition, the comparability issue is presented in the foreword: "... harmonisation of methodology, and in the precision and accuracy of the concepts, definitions, classifications and accounting rules which have to be applied in order to arrive at a consistent, reliable and comparable description of the economies of the Member States." (ESA 1995, page I). Comparability is limited to the results at the macro-level, the data presented by the system of accounts. Data needed at a lower level of aggregation than the full macro-level do not (yet) appear to be important for comparisons of sectors in society or branches of industry life.

Unlike SNA 1993, although the importance of satellite accounts is recognised, they do not merit a separate chapter. It is noted that satellite accounts can be very useful when specific data needs arise, for example for the analysis of the costs and financing of health care (ESA 1995, page 5).

With regard to comparability, it is noteworthy that as in the SNA, comparability is mentioned as being important in the ESA, but in practice not much attention is paid to it. This is all the more peculiar as in the EU the results of the ESA are used to determine the financial contributions to the Union and the distribution of the financial means across member states.

3.4.2 *The development of the ESA*

As a follow-up to ESA 1970, ESA 1995 was adjusted in much the same way as SNA 1993 compared with SNA 1968. ESA 1995 differs in scope and in concepts from ESA 1970. Of course there are some differences which do not result from changes in the SNA.

Although the changes in scope and concepts of ESA 1995 are important for the development of the system, they are less relevant to the present study if they do not influence comparability. The changes in ESA 1995 can be divided into two sections: relevant for comparability and less relevant for comparability.

The following changes can be mentioned as examples of the latter category: the inclusion of balance sheets, a new price-adjusted income concept and use of purchasing power parities (PPP's), the inclusion in production of literary-artistic work, a change in the valuation of insurance services output, chain linking for constant prices, financial lease (see ESA 1995, page 7, 8).

The changes which are relevant for comparability include: households sub-sectoring and a new additional concept of consumption (actual final consumption). ESA 1995 introduces supply and use tables for the first time, already included in SNA 1968. The ESA makes a clear choice for the valuation of produce, i.e. basic prices; SNA still allows producers' prices (see section 3.5.2). The introduction of registration thresholds is specific in the ESA, but not the SNA.

3.5 *SNA, ESA and comparability*

Although all the differences between ESA and SNA are potential sources of problems when comparing the European area with non-European countries, not all these differences are equally important from the point of view of comparability of health care expenditures.

The degree of comparability between the two systems depends on definitions of the terminology, the valuation of transactions and statistical units used for the process, and the classifications used.

3.5.1 *Concepts and definitions*

Although ESA 1995 is consistent with SNA 1993, there are some differences in concepts. Because the European Union stipulated the obligatory character of the ESA and the Netherlands is part of the European area, the description and presentation of the elements used in the comparisons start in the ESA. The concepts used in the ESA are more specific and more precise. For example: ESA concepts list what is included and what is excluded, while SNA describes concepts more in general terms and tries to provide the underlying rationale. This difference in starting point can be detected throughout the description of the two systems.

As an example of the way both ESA and SNA treat a basic element, the definitions of 'institutional unit' will be presented.

Definition in ESA 1995:

The institutional unit is an elementary economic decision-making centre characterised by uniformity of behaviour and decision-making autonomy in the exercise of its principal function. An institutional unit has decision-making autonomy in respect of its principal function and either keeps a complete set of accounts or it would be possible and meaningful (from an economic and legal viewpoint) to compile a complete set of accounts if they were required (ESA 1995, page 19).

Definition in SNA 1993:

Institutional unit: An economic entity that is capable in its own right of owning assets, incurring liabilities and engaging in economic activities and in transactions with other entities (SNA 1993, par. 4.2, page 87).

This example illustrates a general tendency for ESA 95 to recommend more specific concepts and definitions than SNA 1993, which serves more global concepts. The intention of using either set yields similar results.

3.5.2 Valuation and units

Another example of the more specific description of terminology concerns the valuation of transactions. ESA 95 requires that all output⁵⁾ has to be valued at basic prices⁶⁾ (ESA 1995, page 8). SNA recognises the concept of basic prices, but describes this term more globally. Moreover SNA 1993 allows the use of producers' prices⁷⁾ as valuation of output.

In relation to comparability it is clear that as all EU (European Union) countries use ESA 1995, valuation can only be a disturbing factor in the transition period. In comparing only countries using SNA 1993, or countries using SNA 1993 with those using ESA 1995, valuation can be of influence in the results.

The ESA (as well as SNA) is characterised by the use of two kinds of units and two different subdivisions of the economy, both fulfilling different needs: the institutional unit and the local kind-of-activity unit (called establishment in SNA). In the Regulation on Statistical Units of 1993 eight statistical units⁸⁾ are defined, which the member states have to use in "... collecting, transmitting, publishing, and analysing statistical data on the production system, particularly those connected to NACE Rev.1." (European Union 1993, article 2). Because these units are at the centre of the European statistical system, it is important that interpretation of these definitions is the same in all member states.

Interpretation and use of statistical units by the statistical offices in Europe and Eurostat is only one side of the coin. The administrative unit, the company that has to supply all the information is the other side. Information can be supplied directly in the format required by the statistical system, or in a level of detail that the statistical offices can transform into the units needed for the European statistics. Whichever way is chosen, comparability of the end result is not an automatic result.

3.5.3 Classifications

3.5.3.1 Development of classifications in the SNA and the ESA

Classifications are extremely important especially in respect to comparability. Although they do not guarantee comparability, the use of agreed classifications can be helpful at least for improving comparability. Classifications are a tool to create greater comparability. The development of the standard classifications is described

in this section and table 3.1. Classifications are constructed using activities ⁹⁾ as building blocks.

3.5.3.2 SNA and classifications

The first editions of the SNA explicitly used two classifications: the ISIC (International Standard Industrial Classification of All Economic Activities 1949) and the SITC (Standard International Trade Classification 1951). ISIC could be used for the industrial classification of gross domestic product; SITC was the classification of private consumption expenditures. In the first major revision of the SNA in 1968 these two classifications were of course revised and improved. In SNA 1968 ten classifications in all are mentioned and described, ranging from a classification on institutional sectors (SNA 1968, page 78ff), via one for imports and exports of goods and services (SNA 1968, page 117) to one for financial assets and liabilities (page 133). SNA 1993 also makes use of ten classifications, covering the complete area described in the national accounts. Some of the classifications can be traced back to earlier editions, but others are new.

3.5.3.3 ESA and classifications

The ESA uses a (limited) set of classifications, one of which has already been mentioned, the classification of transactions. ESA 1970 uses eight classifications. As in SNA 1968, the basic classifications are present: on industries (NACE), institutional sectors, and final consumption. ESA 1995 uses the same structure of classifications as SNA 1993, which means that the classifications were modified or revised in the European system as well. Only two classifications (COPNI,

Table 3.1
Overview of classifications used in SNA 1993 and ESA 1995

Classification of:	SNA 1993	ESA 1995
Institutional sectors	X	X
Transactions	X	X
Balancing items	X	X
Assets	X	X
Industrial activity	ISIC Rev. 3 ¹⁾	NACE rev. 1 ²⁾
Products	CPC ³⁾	CPA ⁴⁾
Individual consumption by purpose	COICOP ⁵⁾	COICOP ⁵⁾
Collective consumption	COFOG ⁶⁾	COFOG ⁶⁾
Purposes of NPISH ⁷⁾	COPNI ⁸⁾	
Outlays of producers by purpose	COPP ⁹⁾	

¹⁾ ISIC: International Standard Industrial Classification of All Economic Activities.

²⁾ NACE: General Industrial Classification of Economic Activities within the European Communities.

³⁾ CPC: Central Product Classification.

⁴⁾ CPA: Classification of Products by Activity.

⁵⁾ COICOP: Classification of Individual Consumption by Purpose.

⁶⁾ COFOG: Classification of Functions of Government.

⁷⁾ NPISH: Non-Profit Institutions Serving Households.

⁸⁾ COPNI: Classification of the Purposes of the Non-Profit Institutions Serving Households.

⁹⁾ COPP: Classification of Outlays of Producers by Purpose.

Classification of Purposes of Non-Profit Institutions Serving Households and COPP, Classification of Outlays of Producers by Purpose) are not explicitly mentioned and used in ESA 1995.

The contents of classifications have to change to adapt to the economic reality of every day life. In this way classifications offer a framework from which comparisons can start, even at a lower than aggregate macro-economic level.

The development and the growth in number of the classifications used in the SNA and the ESA can be helpful in creating more and better comparability across countries. These classifications are especially created for specific purposes. When all countries use COFOG, for example, expenditures related to the functions of government are better comparable than in case they all use their own classification or no classification at all. The exclusion of COPNI and COPP in the EU classification system is incomprehensible. In the first place, the set of classifications prescribed for use is not complete. Furthermore, there are countries in which NPISH's are more important than private insurance companies in financing health care. However, the introduction of these purpose-oriented classifications does not solve the problem of comparability of certain (economic) phenomena – like health care – across countries. The reason is that these classifications focus on one or more variables, without deleting existing differences between countries concerning these phenomena.

3.6 SNA/ESA and health care statistics

3.6.1 Introduction

Looking at the complete systems (both ESA and SNA) and realising the concepts and definitions used, one might wonder what the problem concerning comparability in general and comparability of health care expenditures can be. Both systems describe the national economy completely and in great detail, using a standard internationally accepted set of definitions and classifications. Health care, as a subdivision of these complete systems, can be looked at from the industry point of view, as producers, or from the government point of view (as producer, consumer and redistribution agent of income and wealth), or from the consumer perspective. For all these purposes the system contains accounts based on uniform sets of concepts, definitions and classifications.

3.6.2 Production

Production ¹⁰⁾ and the distribution of production is one of the most important subjects in the SNA and ESA. For the health care branch and for health care statistics, this item is also of utmost importance.

From the perspective of health care statistics, totals like gross domestic product, gross national product and consumer expenditures, are only relevant as yardsticks, i.e. points of reference for health care expenditures.

GDP and other national aggregates at the macro-level are comparable across nations. This means that the yardstick (usually GDP) used in comparisons at a lower level of aggregation is comparable. Although GDP is the summation of all production unit classes in the economy, and it is comparable across nations, this does not automatically mean that the production classes, i.e. the building blocks of GDP, are comparable. GDP comparability across nations does not imply that ratios created on the basis of these national aggregates are comparable. The disturbing factor in creating ratios of a specific branch, in relation to macro-aggregates is not the denominator, but the numerator: the branch-specific data.

In principle, it does not matter in SNA and ESA where certain activities are performed on the national level, as long as all activities of the economy are recorded somewhere, and no more is recorded than the total of all activities, i.e. double counting of activities is prevented ¹¹⁾. Looking at a section of the total SNA or ESA system the same holds true in a national sense, as long as there are no changes in the classifying conditions in time. However, looking beyond the fence of a single country's economy, differences and similarities can be explained by the contents of the area under consideration.

Because both the SNA and the ESA include all branches of economic activity, and because health care has been a much debated and growing branch of economic activity since the late 1960s, health care is distinguished as a separate category. In the table below the various constructions of the health care branch in the two classifications of industries used in SNA and ESA respectively are presented.

Looking at the classifications used in SNA/ESA concerning the economic process, it is obvious that the health care field has been relatively important since the start of both the systems. Without going into details it can be stated that it is not certain

Table 3.2
ISIC and NACE relevant to health in SNA/ESA

Classifications	ISIC/NACE			
	Division	Major group	Group	Title
SNA 53	8	82	822	Medical and Health
SNA 60	8	82	822	Medical and Health
SNA 64	8	82	822	Medical and Health
SNA 68	9	93	933	Medical, dental, other health and veterinary services
SNA 93	8	85	851	Health care
ESA 70	9	95		Services of public health and veterinary services
ESA 79	9	95		Services of health
ESA 95	N	85	851	Services of health

SNA: System of National Accounts.

ESA: European System of Accounts.

ISIC: International Standard Industrial Classification of All Economic Activities.

NACE: Statistical Classification of Economic Activities in the European Community.

whether the contents of the major (ISIC or NACE) groups are comparable, or whether they have remained the same or comparable in the course of time. It might well be the case that the use of the data as presented in the national accounts of any country shows remarkable changes caused by classification changes. On the other hand, the changes as presented in table 3.2 might well be limited to the names and the contents of the major groups, and while the groups themselves have remained the same.

It cannot be stressed too often that classifications (e.g. ISIC and NACE) are not created with the purpose of cross-country comparability of statistical results (e.g. expenditure). Classifications are made to be able to categorise all items (units, activities, etc.) in a unique way, based on a (limited) set of criteria to be applied. Even when the criteria are the same across countries, the data collected and presented according to the classified units, are not by definition comparable. The use of classifications at a given level (e.g. ISIC 85 or NACE 85), however, can help to limit comparability problems by providing insight into the location of activities and units of production. The more detailed the level in the classification to be compared across countries, the greater the problems encountered and the greater the importance of distinct (sets of) activities become. Classifications can be adjusted as a consequence of this large amount of information on the location of activities. If this is the case one should be aware of the fact that if these changes are not carried back in time, time series become distorted.

3.6.3 *Consumption, expenditures and financing*

“The term consumption on its own can be ambiguous and misleading. Sometimes it is used by economists to refer to consumption expenditures, sometimes to acquisitions of consumption goods and services and sometimes to the physical use of goods and services for the direct satisfaction of human needs or wants.” (SNA 1993, page 213).

The principle that all relations are basically linked to sectors compounds the problem of ambiguity in the terminology. For consumption this means that the household sector is involved as well as the government sector. The classification of consumption according to the sector of expenditure is not the only one users of SNA data are interested in. Which units (individuals, groups of individuals or the population at large) actually benefit from the goods and services is also of interest. This means that a distinction of consumption into individual consumption and collective consumption is necessary.

Government, however, is not only a sector in the SNA related to consumption (expenditure) and production. Government also has a role in redistributing income and wealth. Government can redistribute income and wealth across the population using goods (including financial goods) and services. For the services branches the redistributive role of government is linked to financing. The primary financing unit

is not always the ultimate bearer of the expense (e.g. through reimbursement strategies and third party payer systems).

These topics (terminology, organisational influences, government roles) have been subject of study in recent decades.

The way the goods and services are paid for primarily and ultimately, influences the levels of consumption expenditures of the two sectors involved (government and households). Especially, when the systems of financing of the health care branch are changed (from privately financed systems to publicly tax based funded systems, or vice versa), breaks in consumption expenditures series will result. "Strict application of the SNA [and ESA] rules would have entailed ... a break in their time series on final consumption expenditure of households and on government consumption expenditure, since the health expenditure ... passed from one category to another at the moment the new health scheme was applied." (Pêtre 1983, page 26). The breakdown of final¹²⁾ consumption expenditure between households and government based on the degree and methods of government interference makes international comparison of the results almost meaningless.

From the viewpoint of international comparability on health care expenditures of households, two solutions are open. Either "... it would have been desirable to consider ... the supply of medical goods and services as social benefits in kind and the corresponding goods and services as part of final consumption of households ..." (Pêtre 1983, page 14) or alternatively treating "... unrequited transfers for specific functions as government consumption expenditure." (UN 1986, page 17). Neither solution is politically or statistically desirable. However, the second solution has become the rule in SNA 1993 and ESA 1995.

A less disruptive way of solving the problem (according to Pêtre) is to divide the subject into separate but, of course, connected items. The terminology of market/non-market is related to production and should only be used in connection with the production process and the valuation of the goods and services produced. For the way the expenditures are paid for, the primary and ultimate payer distinction could be introduced, leaving individual and collective as terminology for consumption. Thus every part of the economic process described (production, consumption and financing) has its own set of distinctions.

The other aspect, especially the paying aspects of health care services (and services in general) is not yet completely solved. The redistributive role of government is recognised in SNA/ESA but not yet completely resolved either.

For both items a first step towards the solution of the problem is the introduction of the notions actual consumption and consumption expenditures. After extensive debate these two distinct terms were introduced in SNA 1993 and ESA 1995 to delete this ambiguity on consumption. The term consumption expenditure is used to identify the institutional units that control and effect the expenditure. The term actual consumption is used to identify the units that actually benefit from the expenditures.

In conclusion it may be stated that some of the problems concerning the contents of the terms used have been solved, especially those relating to consumption by institutional sectors. However, certainly not all of the problems have been solved, and the solutions are not definite.

3.7 Comparability and conclusions

Although international comparability is one of the criteria in the development of an SNA, and comparability is one of the major features of quality, in practice not very much attention is paid to the problems surrounding comparability of the results.

3.7.1 Comparability and SNA/ESA

National accounts supply information on the economy in an aggregated sense, allowing comparisons with the past (given classifications and definitions).

Dernburg and McDougall note that the usefulness of GNP (Gross National Product) is applicable to any country at any point in time, but that "... its usefulness is certainly greater when applied to highly developed countries during recent periods. The farther one gets from developed market economies the less useful is the rule for distinguishing economic activity from the general activity of life. The rule means that many of the activities carried on within households are excluded from the measurement of the nation's product." (Dernburg and McDougall 1972, page 55). Implicitly one of the problems relating to comparability is mentioned here, the boundary problem. This problem, however, is not confined to household production (see chapter 2, section 2.3.3. production boundary) but crucial in comparisons across countries.

Although it can hardly be concluded from the publications on national accounts (except for some small remarks scattered throughout the various SNA and ESA publications), national accounts experts are sure that "...international comparability of macro-economic data can be considered as the main goal of international guidelines." (NOS 2000a, page 24). In this respect the introduction of both the new SNA 1993 and the new ESA 1995 are a major step forward.

The use of national accounting data, even the administrative use of these data, may have some serious drawbacks, but it is certain that it served to boost the progress made in the field of international comparability of economic data at a macro-level.

In the use of branch information comparability (and its lack) across borders and in time is eminent. Although the degree of quantitative precision as mentioned on page 2 in SNA 1953 now has a completely different meaning, it can still be a very limiting factor in international comparisons. In 1953 the absolute lack of data would be the problem. Now, at the beginning of the 21st century, the level of detail in the most detailed parts of the system is more likely to be the limiting factor.

One of the larger problems concerning the information available in national accounts on the sub-national level was the rigidity of the system itself. National accounts do not provide the information needed at management and policy level. Some flexibility was introduced in the 1993 version; although this might solve problems for information needed for national policy issues, it does imply the possibility of deviating from the agreed standards, which results in less comparability of data supplied across countries.

On the other hand we have to remember that the system of national accounts was not developed for the purpose of collecting internationally comparable national accounts data; the SNA was developed for national purposes and to be "... sufficiently comprehensive, that individual countries, whatever their economic structures, ... can select from within it those parts which are considered to be most relevant and useful to implement ..." (SNA1993, page 8). It is because the SNA has become a standard system used throughout the world with little or no modification, that (international) organisations use it.

International organisations use the SNA to compare major aggregates, to compare the performance of economies with each other, to influence popular as well as political judgements about the success of economic programmes, and for econometric analysis (time series and cross section). International and supranational organisations (e.g. the Worldbank, IMF and the European Commission) use the data to determine the eligibility for loans, aid or other funds and to determine the conditions for these loans and aid programmes.

Although national accounting data seem to be undisputed at the macro-economic level, an investigation by the EU Court of Auditors showed that the quality and reliability, especially comparability, of macro-economic aggregates was unsatisfactory (see Eurostat 1999b, page 2). This report also states that reducing the errors in national accounts is difficult, has high demands in terms of information storage and processing and takes up a lot of time and resources. In the first place the great number of source statistics used in national accounts implies that accuracy assessment is required as input information for the quality assessment of the national accounts statistics. In the second place there is as yet no consensus on how the various error types (related to input and compilation process of national accounts) can be aggregated into a single error rate for GDP.

If the data are to be compared at a lower level of aggregation than the national totals, comparability problems increase. At the national level only one boundary, the production boundary, has to be taken into consideration; at the lower, industry level, every branch of industries has its own boundary. The boundaries of the same branches of industry are not identical in all countries. The most important disturbing source is the way the economy is organised; this organisation, which levels out at the national level, can be a serious source of disturbance at, e.g. ISIC or NACE 2, 3, or 4 digit level. The higher the degree of detail in classification used in the

comparisons, the larger the disturbances and the larger the problem of comparability will be.

So, although classifications do not completely solve comparability problems, their use can assist in the enlargement of comparability across countries.

Besides production, consumption and consumption expenditures are main topics in both SNA and ESA. Although several areas of ambiguity concerning the terminology have been erased in SNA 1993 and ESA 1995, the problems regarding consumption-related topics have not been completely solved. For cross-country comparisons the proposals made by Pêtre (on the distinctions between production, consumption and financing) might still prove to be very useful.

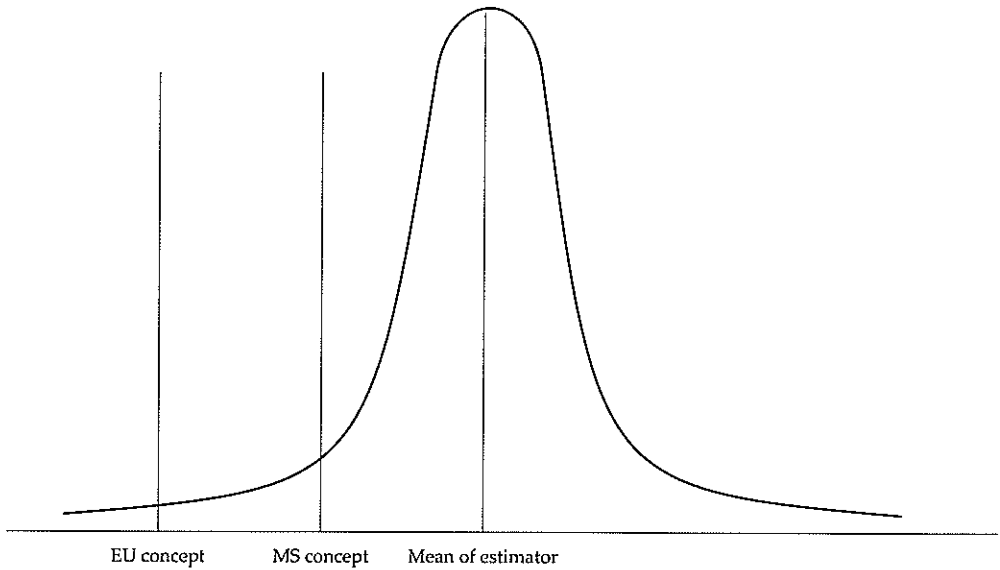
National arrangements and resulting national differences in sub-national boundaries can be tackled by using the allowed flexibility of national accounts (by means of satellite accounts) or by creating a system of functional accounts. Satellite accounts and functional accounts have the same generic defects, because they (usually) use the same classifications and the same sectoring. Neither the use of satellite nor functional accounts can solve the comparability problem across countries for production groups or production classes in the economy, however. Very detailed meta-information – for example on the level and the kind or the combinations of activities performed in the production process – is needed, to decide on and improve the level of comparability of the accounts under study.

3.7.2 Comparability and standards

This factor, the consequences of the national arrangements, is hardly ever mentioned, and certainly not in the publications on SNA and ESA. It can be seen in the differences in both definitions and classifications resulting from national interpretations of the international systems. In a theoretical approach to the measurement of comparability (see Figure 3.2) the difference between national and international concepts are visualised in the EU concept and the member states' concept (Eurostat 1999a, page 15). In this figure the member state concept (MS concept) and the European Union concept (EU concept) are two different interpretations of the system relating to one kind of statistics. The mean of estimation represents the results of one or a limited set of surveys on the subject.

Ultimately this difference could lead to two different datasets for one country, one based on local national needs and one based on the international standards used in international and supranational organisations. Comparability across countries can be boosted if these dual systems are eliminated or reduced to a minimum, i.e. if national systems are adapted to the international standards. The ideas incorporated in this theoretical approach are valid for both national accounts aggregates and data at branch level (e.g. health care).

3.2 Theoretical approach to comparability



Health care is an outstanding example of a branch in which national rules and regulations, national sentiments and organisations play an eminent role and cause the existence of different systems. The European Commission does not consider harmonisation of the various European health care systems as an explicit target area. Classifications and definitions, even those based on the agreed international systems, will incorporate the local national peculiarities and influence the results. Health care statisticians who want to comply with internationally agreed rules and regulations nearly always have to convert their national data to an agreed international dataset. Although harmonisation is not an explicit target, it has already resulted in narrowing the gaps between data from national health care systems, and will almost certainly continue to do so in the future.

3.8 *Final remarks*

With regard to the questions posed at the beginning of this chapter the following conclusions can be drawn.

Both systems described and discussed in this chapter, the SNA and the ESA, set out criteria and starting points that statistical systems are supposed to use. Comparability, however important it is deemed to be, is only mentioned sporadically. Complete and comprehensive coverage of all economic activities in statistical systems is the most important issue.

In practice little attention is paid to cross-country comparability, not at the aggregate level and especially not at a lower than aggregate macro-economic level.

Some progress in this respect can be observed in the course of the development of both systems with the introduction of instruments to increase flexibility, such as satellite accounts (only in SNA) and functional classifications. These items were introduced primarily to provide more insight and to make the systems more useful for analyses at a lower than aggregate level. They do not solve the comparability problem, however, as the contents of the phenomena to be compared remain different. Although classifications, certainly internationally agreed classifications, are a necessary condition, the existence and even use of these classifications do not guarantee cross-country comparability.

The conclusion that can be drawn from the perspective and the systematic approaches in SNA and ESA in relation to comparability, is that they are promising when dealing with macro-aggregates but less useful at lower levels of aggregation, because of differences in contents of the areas to be compared. Health care data (in some cases including social care data) based on the various national accounts sources (national where possible, and international) for EU and non-EU countries will be presented and discussed in Chapter 9.

But even in national aggregates at the macro-level differences may result from different interpretations (e.g. MS concept versus EU concept). In principle all these differences can be documented and the concepts can be described. Whether, and to what extent the resulting differences can be quantified is another question altogether, a question that probably does not have a very high priority either in national statistical offices or in Eurostat.

Notes:

- 1) In SNA and ESA the following institutional sectors are distinguished: Non-financial corporations, Financial corporations, General government, Households, Non-profit institutions serving households and Rest of the world.
- 2) A transaction is an economic flow that is an interaction between institutional units by mutual agreement, or an action within an institutional unit that is useful to treat as a transaction, often because the unit is operating in two different capacities. (ESA 1995, page 9).
- 3) To be completely accurate, the OEEC (Organisation for European Economic Co-operation) published guidelines at the end of the 1940s, which the SNA 1953 extended to cover the whole world.
- 4) Technically European integration started with the creation of the Benelux (1948) and the European Coal and Steel Community (1951).

- 5) Output (see ESA 1995, page 34) consists only of those goods and services that are produced within a local KAU (Kind of Activity Unit, establishment) that become available for use outside that local KAU (establishment).
- 6) The basic price (see ESA 1995, page 44) is the price receivable by the producers from the purchaser for a unit of a good or service produced as output minus any tax payable on that unit as a consequence of its production or sale (i.e. taxes on products) plus any subsidy receivable on that unit as a consequence of its production or sale (i.e. subsidies on products). It excludes any transport charges invoiced separately by the producer. It includes any transport margins charged by the producer on the same invoice, even when they are included as a separate item on the invoice.
- 7) The producers' price is the amount receivable by the producer from the purchaser for a unit of a good or service produced as output minus any VAT, or similar deductible tax, invoiced to the purchaser. It excludes any transport charges invoiced separately by the producer (SNA 1993, page 151).
- 8) These eight statistical units are: the enterprise, the institutional unit, the enterprise group, the kind-of-activity unit (KAU), the unit of homogeneous production (UHP), the local unit, the local kind-of-activity unit (local KAU) and the local unit of homogeneous production (local UHP).
- 9) An activity is said to take place when resources such as equipment, labour, manufacturing techniques, information networks or products are combined, leading to the creation of specific goods or services. An activity is characterised by an input of products (goods or services), a production process and an output of products. (NACE rev. 1, Introduction, page 9).
- 10) Although no separate definition of production value is presented, neither in SNA nor in ESA, it is possible to derive a description for this variable (see ESA 1995, pages 37, 13 and 39): production value is the value of the goods and services produced measured at market prices seen from a supply or supplier's point of view. Market prices for market output, output produced for own final use and total output of market producers and producers for own final use are basic prices (see footnote 6), while the total output of other non-market producers is valued from the cost side.
The total cost of the production (see ESA 1995, page 45) consists of the sum of intermediate consumption, compensation of employees, consumption of fixed capital and other taxes on production less other subsidies on production.
- 11) Double counting of financing is not necessarily prevented, due to the intricate web of cross-subsidisation. Only if the recipients of the financial flows are fully known, can double counting be prevented.
- 12) Final meaning that these goods and services are not reused in any production or redistribution process, but serve the purpose of satisfying needs.

4. *Satellite accounts*

4.0 *Questions*

Is there any international agreement on the construction and use of satellite accounts?

What are the basic criteria and starting points in satellite accounts?

How are these criteria and starting points relevant to comparability at branch level?

Does the construction of satellite accounts reduce the shortcomings concerning comparability in the core system of the national accounts, and if so to what extent?

4.1 *Introduction*

The previous chapter dealt with the possibilities of the central framework of the System of National Accounts, SNA (and European System of Accounts, ESA). The SNA – an integrated accounting structure with a coherent set of concepts – is exhaustive and consistent within the boundaries of economic activities. The disadvantage of the SNA, and also the ESA, is in its limitations. The SNA gives each unit, transaction, product and purpose a unique place in the classifications. The various uses made of the SNA (economic analysis, forecasting, policy), call for different approaches. Because these approaches are not homogeneous (different economic theories, various forecasting techniques and a host of policy issues) and not constant through time, they all have their own specific data requirements.

Statistical agencies have to deviate from the rigid framework of the SNA to make data more suitable for politicians, scientists and other users respectively. Satellite accounts may be one way to supply the data needed for these different purposes.

In this chapter the possibilities and limitations of satellite accounts will be discussed in view of their usability for international comparison of the production value ¹⁾, consumption and financing of health and social care.

4.2 *The history of satellite accounts*

After the Second World War the construction of a completely developed set of national accounts created a quality leap forward in the whole area of general, monetary value statistics. In later years this development was followed by integrated input/output (I/O) tables and financial accounts. The combination of a coherent set of accounts, including financial accounts and the availability of an input/output table, generates a solid base to create satellite accounts to the national accounts.

The users of the statistics soon began to demand that statisticians extend these qualities of coherence and up-to-dateness existing in national accounts to a more detailed framework covering a more specific field of information.

At the end of the 1960's and the beginning of the 1970's there was again loud criticism of the system of national accounts and its most important aggregates. First of all national accounts are limited to economic data. This restriction, more precisely to evaluation in monetary units "... would make politicians concentrate their decisions and measures primarily on those areas for which ... hard data are detectable." (Haslinger 1988, page 60). Areas for which no 'hard data' are available are much harder to evaluate.

Secondly national accounts deal with national aggregates at the macro-level, while politicians, although interested in aggregated data, are much more interested in data on areas of social concern. Policy relevant areas do not always correspond with classifications and definitions used in the national accounts. Erroneous trends, for example in health care and other areas of social concern could not be avoided by timely intervention, because of conflicting interests of national accounts and policy areas. The data supplied by national accounts on areas of social concern do not correspond with the demand for data at the policy level, which usually concerns the meso-level (i.e. the branch) and the micro level (i.e. the type of provider or financing unit). Certain general information about the health service can be derived from national accounts, particularly where state intervention is concerned, but they cannot provide a more detailed view of the economic dependency of the health service in relation to other economic areas; their objective is too generally economic.

A third point of criticism was that national accounts were primarily devised in a period when emphasis was put on the production of goods, one element not mentioned by many authors. The production of services was limited in volume. Nowadays, the services industry, both commercial and non-commercial services, is much more important. Services are not treated differently from goods in the SNA. Foulon (Foulon 1982, page 46) touches on this by stating that "... this division [between goods and services] has not had any important effect on basic economic analysis since goods and services are equivalent from the point of view of general equilibrium, the theory of international trade or income distribution." However, services are not just intangible goods which can be produced; they cannot be put in

stock. Many services have some aspects of investment along with consumption, depending on the view taken. Health care expenditure, and more specifically medical care expenditure, can be seen as investment in human capital. The same is true for education expenditure.

On top of that, for a large section of all services there is no real market mechanism where demand and supply fix prices. In many cases the government determines the 'prices', and in some cases also the amount of services to be supplied. Following the rules laid down in SNA and ESA, market prices ²⁾ should preferably be used to determine the value of government controlled production. If market prices for analogous services are not available, valuation should be based on production costs.

The fourth and last point of criticism is more general by nature. After the revision of the SNA (SNA 1968) and the construction of the first ESA (ESA 1970), the user demand for co-ordinated statistical results was felt more strongly and even became critical in all countries. The relative underdevelopment of social statistics might be seen as an indirect consequence of the priority given to the development of the SNA.

Basically two different solutions are possible to incorporate a given issue in national accounts: complete integration in the framework of SNA, or the creation of a specific system. Complete integration has the advantage of a coherent system, but puts in place restricting structures and methods. The creation of a separate system eliminates these restrictions, but introduces a lack of coherence with the central system. In France a middle course was chosen, in an attempt to combine the advantages of the two solutions, while at the same time trying to minimise the disadvantages (see Laganier 1982, page 55, 56).

The idea of satellite accounts was put forward in France as early as 1967, largely prompted by the shortcomings mentioned above. Vanoli was one of the first researchers to use this term in a paper (see Vanoli 1969, page 183). The idea of satellites to the national accounting framework was based on the French ³⁾ experience with specialised accounts in the 1960's and on the analysis of the instrumental needs of planners. French national accountants were convinced of the necessity of providing a certain degree of flexibility and thus autonomy for statistics, to develop and specify national accounts in specific fields, like health care, housing and education.

In France the first experimental work explicitly dealing with a satellite account was started in 1968, namely on housing. In 1969 the development of various sets of specific instruments around a central framework of national accounts integration was proposed and agreed. These sets showed various degrees of integration in coherence with each other and with the central framework. Several instruments were designed and satellite accounts centred on specific fields of public policy concerns. The French satellite accounts on research, health, education and social protection reached the stage of publication at the end of the 1970's. The satellite accounts on health (*Les comptes de la Santé, Méthodes et Séries, 1950-1977*) were

published by INSEE (Institut National de la Statistique et des Études Économiques) in collaboration with CREDOC (Centre de Recherche pour l'Étude et l'Observation des Conditions de Vie) in 1979.

Outside France there was less enthusiasm for the creation of satellite accounts to the national accounts. Only a few other statistical offices (e.g. Germany, Norway and the USA) addressed the issue of satellite accounts before the publication of SNA 1993. Of the international organisations, the Organisation for Economic Co-operation and Development (OECD) ⁴⁾ first mentioned the concept of satellite accounting in relation to tourism in 1985. Although they can solve some problems caused by the rigidity of national accounts, satellite accounts are still not a full-fledged system with a stable set of own definitions and own classifications.

The theoretical work on satellite accounts went on in the academic field and in statistical organisations, resulting in a complete chapter (chapter XXI) in SNA1993. But ESA 1995 makes no mention of satellite accounts and the use that could be made of them.

4.3 *Scope of satellite accounts*

All experts limit the construction of satellite accounts to fields of social concern (an extension of Vanoli's original idea of "... functional analysis ... of public expenditure ...", see Vanoli 1969, page 183). Fields of social concern, however, is a vague term; the boundaries are not clear or restricted. If fields of social concern are interpreted as all areas about which society would like to have coherent information, then almost no topic is excluded. On the other hand, social concern can be limited to fields of collective concern, regarding public expenditure. A third possibility is to limit social concern to areas of social life and social care, like health, education, and wellbeing, generally areas in which the government intervenes actively because there is no market for public goods, or because external effects are important ⁵⁾. France uses a broader concept of social concern than elsewhere in the EU (European Union), resulting in the construction of satellite accounts for numerous areas, including some not defined by social concern (e.g. transport, IT, R&D, agriculture).

Pommier was the first to distinguish several families of satellite accounts, "... some based on economic activities, ... others aiming at population needs ... others, ... in relation to alcoholism or environment, ... would be rather different from the first two families." (Pommier 1981, page 376). In 1988 Teillet limited the number of families for which satellite accounts can be made to three (Teillet, 1988, page 415–417): the first set consists of the compilation of economic satellite accounts, especially useful for activities which are not presented in a clear and exhaustive manner in the central framework, for example agriculture or trade; the second set is the 'generalised

functional analyses', which were the first analyses to be presented under this name, like tourism and health; and the last set of accounts to be constructed concerns activities of sole public authorities, like public aid and alcoholism.

Teillet puts generalised functional analysis between quotes. Functional analysis probably refers back to Vanoli's original idea of functions of the public domain. 'Generalised' could refer to an extension of this terminology, meaning that the area is not limited to public expenditures (but relates to other variables and/or more overall population needs). On the other hand 'generalised' may also refer to the central framework of the national accounts, because "... satellite accounts combine an extension of the kind of activity and product analysis and a generalisation of the purpose approach." (SNA 1993, par. 21.51, page 494).

Schäfer and Stahmer (Schäfer and Stahmer 1990, page 169, 170) oppose such a broad interpretation of the subject field of satellite accounts, among other things because the concepts used in the three families of satellite accounts are not the same. They propose that the term satellite accounts be limited to generalised functional accounts as presented in the French system. Although in SNA 1993 the emphasis is on functionally oriented satellite accounts (SNA 1993, chapter XXI, section C, page 494–508), the other families are not excluded.

In my opinion, it is not necessary to exclude detailed analysis of economic activities from satellite accounting. Both detailed analysis of economic activities and the generalised functional analysis and accounts start within the activity classification of the NACE (Statistical Classification of Economic Activities in the European Community) and the ISIC (International Standard Industrial Classification of All Economic Activities), and use statistical and economic tools for the analysis of the area. The same kind of reasoning can be applied to the exclusion of the analysis of areas like public aid or alcoholism. These areas cannot be identified with a complete production class, or parts thereof, in the ISIC or NACE classification, but that is true for tourism as well. As long as it is possible to select parts of production classes for detailed analysis, satellite accounts can be produced. Neither the approach of the accounts (production of goods and services or purpose of economic transactions), nor the subject (economic activities or activities of social concern, population needs) should on their own be decisive elements. What should be decisive for being a satellite account is the seamless link to SNA and the use of agreed methods in the construction of the accounts.

Tables on production, financing and consumption are the minimum basis to be able to answer the most important questions on who produces what for whom, who pays for it and, lastly, who benefits from it. In practice the construction of a satellite account will not be performed by extracting data from existing accounts (top down), but rather by starting from the same basic material used for the construction of the SNA accounts (bottom up).

4.4 *Satellite accounts: definitions and descriptions*

In the 1960's and 1970's criticism on the SNA became increasingly apparent, a criticism followed by constructive efforts to create satellite accounts. In the 1980's, however, there was still no agreement on the contents of concepts and structure of satellite accounts. There was no international standardisation and it was clear that not all arithmetic endeavours that could be brought into a relationship with the SNA could be considered as satellite accounts.

4.4.1 *Satellite accounts: general description, features and characteristics*

Various authors described satellite accounts in the 1970's and 1980's. Pommier (Pommier 1981, page 376) describes a satellite account as "... an information gathering framework, open to change, in a sector of social concern. It presents two basic characteristics: it brings together monetary and non-monetary data in the same presentation and it unites a detailed analysis of the economy of a sector to overall analysis."

Haslinger (Haslinger 1986, page 66) describes a satellite account as "... a consistent system of monetary and non-monetary measured variables ... that can provide sufficiently exactly in detail and completely the processes and situations and changes thereof, ... referring or connecting to an important social item. The monetary measured values are to be linked to the central framework."

In chapter XXI of the SNA 1993 the characteristics of satellite accounts, as presented by the various authors, are summed up in the following description:

Satellite accounts provide additional information on topics of social concern, "... using complementary or alternative concepts and classifications to introduce additional dimensions to the central framework ..." and provide extensive "... coverage of costs and benefits of human activities ...", while linking physical sources to monetary data (SNA 1993, par. 31.4, page 489). These complementary or additional concepts are not to diverge drastically from the concepts used in the central framework, to retain the possibilities of linking.

As already mentioned, one basic feature of a satellite account is that it is linked to the central framework of the national accounts. Although all satellite accounts experts agree on this, they differ with respect to the type and strength of the links. For some it is enough that some classifications developed for the national accounts are used, for others, using deviating classifications to incorporate the results in the central framework is sufficient. The majority, however, are of the opinion that aggregates created by a satellite account should fit seamlessly into the central framework. This does not mean that a satellite account is confined to the same NACE or ISIC classes defined in the SNA or ESA, because "... the satellite account approach includes the liberty to cross the borders of the sector in question and the rest of the economy ..." (Hansen 1996, page 3). The only real limitation is that the aggregates of

corresponding NACE or ISIC groups add up to the aggregates as presented in the national accounts, and thus satellite accounts are formally consistent with the national accounts.

Linking detailed analysis and data to national accounts does not call for a separate system. As Stahmer mentions in his article, in the relation between national and satellite accounts: "Information ... could surely be got by a further breakdown ... of the national accounts. Then the possibility would be available for the users to extract the requested data ... as from a component system. The methodology to subdivide existing accounts further ... would not change the concepts fundamentally. The consequence however would be that the system ... became unclear ... [and] it might be difficult and laboriously ... to find the data." (Stahmer 1988, page 11).

A second feature of satellite accounts is that they incorporate the possibilities to include non-monetary information. Non-monetary values are not limited to inputs or outputs of the production process, production factors or intermediate products, but can also contain indicators for measurement of the result of the interventions, and breakdowns of beneficiaries by all kinds of characteristics.

A third characteristic of satellite accounts is the level of freedom in applying, for example, deviating or additional concepts and classifications. It is possible to change the production boundary of the national accounts by including household activities of private households. Of course if the production boundary is enlarged the consequences for the complete system of accounts has to be traced (production, income, income distribution, consumption, investment).

Lastly, all researchers place most emphasis on expenditure connected to the field in study (how expenditure is used, how it is financed and who benefits). Expenditure is usually presented as a special case of the general SNA definition: a way to express the exchange of goods and services between a purchaser and a seller, from the paying or purchaser's point of view. In satellite accounts and other forms of health accounts the expenditure is usually interpreted as a financial representation of the efforts of society for having and maintaining a health care system.

4.4.2 *Satellite accounts: framework and definitions*

Combining the various elements mentioned above, i.e. linking to national accounts and the subject of social concern, in practice the subjects of satellite accounting may well be limited to generalised functional analyses. A satellite framework or account (escaping from the institutional constraint of the central framework) is by hypothesis more functional.

Chapter XXI of SNA 1993 provides the following 'starting points' additional to the general criteria of SNA (SNA 1993, par. 21.5, page 489). Satellite accounts:

- are linked to the central framework;
- are linked to the specific information system of the field of study;
- aim to integrate financial and non-monetary data.

By limiting satellite accounts to topics linked to SNA, the analysis of the field of study in a macro-economic context is easier, and the relations between various separate satellite accounts in the national economy are determined.

If these guidelines are followed internationally it makes it possible to compare satellite accounts on one topic between countries. The advocates of satellite accounts sought flexibility within SNA to be able to produce data on a certain topic (like health care) to be used in policy analysis and the implementation of policy measures. In practice the construction of satellite accounts never achieved this goal. Opposite the advantage of better comparability, there is the disadvantage that politicians or managers in areas of social concern may not find the data they need in satellite accounts constructed in this way.

4.4.3 *Description of satellite accounts on health care*

Health care is a prominent branch for which the construction of satellite accounts can be very promising. The description of a satellite account on health care is derived from and linked to the more general descriptions of satellite accounts mentioned above.

To make the link with SNA or ESA more prominent, first of all the term health in the INSEE description should be replaced by health care. Health care is an economic activity that can be measured and defined even within the boundaries of the SNA. Health is not an activity; it is a status (see chapter 2). Secondly, satellite accounts on health care should be limited to production, consumption (and investment) and financing firmly linked to SNA, although ultimately all the accounts needed for the complete system of SNA should be produced for a satellite account as well.

A satellite account on health is described by Essig and Reich as "... an account that describes the economical meaning of the health services, ... globally, clearly, sufficiently arranged and quantitatively. From this follows that substantial activities to health services are to be included. Typical of a Satellite System is that the calculation is as far as possible in agreement with the general SNA." (Essig and Reich 1988, page 87).

INSEE (INSEE 1979, page 9), one of the founding fathers of satellite health accounting, describes the purpose of satellite accounts on health as "... to analyse the various characteristic activities of the field of health, from the viewpoint of financing and of the organisation of the production (functional criteria); describe and analyse the allocation of a certain class of goods and services; [and] to measure the total effort to finance health, agreed by the community."

Keeping in mind the characteristics mentioned above, for me a satellite account can be defined as a flexible system describing economic relations (i.e. related to expenditure, production value and financing) of a limited area of concern (e.g. health care), linked at a certain level of aggregation (in terms of ISIC/NACE) to the

national accounts, providing possibilities to include non-monetary information and to deviate from SNA if the need arises.

4.5 *Contents, scope and construction of satellite accounts on health care*

4.5.1 *Introduction*

Before going into the contents and scope of satellite health care accounting, it may be concluded that there is international agreement on satellite accounts on one point, i.e. linkage to the central framework. However, there are various possibilities for the exact contents and conditions of these links to SNA. Even the chapter on satellite accounting in SNA 1993 offers no definite solution; it gives two options, one offering more flexibility within the central framework, and another in which all concepts, classifications and boundaries could be changed (SNA 1993, par. 21.45, page 493). There is also international agreement on the topics to be dealt with in satellite accounts, i.e. areas of social concern, areas of population needs. However, nowhere are the contents of these terms specified.

The following assumptions are made concerning satellite accounts on health:

- health is to be read as health care, as described in chapter 2, and is in agreement with the general rules of SNA;
- the links to SNA mean that at the aggregate level of the satellite account, the financial data are to be traced back into the SNA at the appropriate level;
- links are to be established at least for the production and the consumption blocks.

4.5.2 *Contents, scope and construction*

By combining the INSEE description of health and the characteristics of satellite accounts in SNA, and supplementing these by some features of the description of health care in chapter 2, the field of the satellite account on health care can be circumscribed as follows:

- characteristic health care activities;
- production and use of all (characteristic as well as connected) health care goods and services;
- the financing of health care (both from public and private funding).

The first and basic part for the construction of a satellite account is the determination of the characteristic activities and goods and services specific to the field. In our case then, the first step is to determine the goods and services that are specific to the branch of health care, and to separate the characteristic from the connected goods and services. Characteristic goods and services are typical for the field of health care and are described in great detail. Connected goods are not typical by nature, or are classified in broader categories of products, but information on these products is of interest because they are covered by the concept of expenditure.

The second part concerns production. Characteristic goods and services originate in characteristic activities and are produced by characteristic producers. Characteristic production can be performed as a principal activity, a secondary activity and even as an ancillary activity in the production process. Characteristic production by characteristic producers is studied in detail.

The third part of the system of satellite accounts is the financing section. Especially in health care, the users do not always bear the expenses; third party payers are common in this branch. Analysis of the ultimate bearer of the expense is central in the core framework.

According to the SNA, characteristic goods and services of the health care branch may consist of health services, public administration in health, health education and research and development in health. In the SNA for the health care branch, patient transport, pharmaceuticals and other medical goods are presented as connected goods and services. So the SNA defines no rigorous boundary of the health care area in terms of which goods and services are characteristic.

Some differences between SNA and other research can be noted, even when health is read and interpreted as health care. One important discussion issue is the boundary of the activities to be included. The core business of the health care branch is the execution of health activities, leading to health services. In SNA this term is not explained (health care is just used as an example of satellite accounting in chapter XXI of the SNA 1993). In the French publication (INSEE 1979) no distinction is made between characteristic and connected services; all medical, dental and paramedical activities, and activities in the area of collective prevention are mentioned as characteristic activities. Agreement on the contents of the set of characteristic activities and characteristic goods and services is necessary for international comparability.

As some statistical agencies do not agree with the inclusion of activities on education and research, and leave them outside the satellite account on health care, there is no agreement between various agencies on the boundary of health care. These activities (e.g. education and R&D) can be assumed to be part of other satellite accounts, and their inclusion could lead to double counting, if they are not separately shown in all the accounts created.

Some experts want to include all production units, even those units with a very small amount of characteristic goods and services in ancillary health care activities, compared with total output. Others want to limit the area and exclude all units offering only ancillary health care services.

Some statisticians treat all retail trade as production of connected goods and services. Although not important in the construction of the production account, the level of detailed analysis is influenced by the determination of retail trade as characteristic or connected goods.

Another point of discussion is the use of the expression public administration in the SNA. Most experts do not limit administration of the health care activities to public

administration. All administration, public as well as private is included. So here too, there is no agreement on the contents to be included in health care.

One of the basic features of SNA is that production can be either market or non-market. The difference between the two is related to the valuation of the production. "Market output, output produced for own final use and total output of market producers and producers for own final use are valued at basic prices, while the total output of other non-market producers is valued from the cost side." (ESA 1995, page 39). If a satellite analysis compares countries with different production structures for the same area of social concern, these comparisons might be heavily distorted by the use of the different valuation techniques.

Until now the focus has been on the production of health care. Production¹ is a basic characteristic in SNA. In areas of social concern, i.e. for managers and policy departments, production is not of prime importance, but expenditure⁶⁾ by society on maintaining and operating a health care system is. In national accounts terminology, however, production value is not the same as expenditures⁷⁾.

As 'third party payer' constructions are common in the health care branch, it is necessary to divide the financing of health care goods and services into two accounts: one for the allocation of financial means to the producers (initial financing) and one on the ultimate financing, the ultimately paying organisations.

Basically, consumption of health care services poses no real problem. For managers and policy departments health care services are either consumed by households (national or abroad) or by the population at large. In national accounts consumption is divided into actual consumption and consumption expenditures and by sector (households, non-profit institutions serving households, and government). These same distinctions need to be followed in a satellite accounting framework.

4.6 *Satellite accounts on health care and comparability*

A satellite account is not just a detailed description of a branch in the national accounts; the added value of these accounts is in supplementing this information by other financial and non-financial data. The financial data can be scattered throughout the national accounts, or may even not be included in them. The non-financial data stem from various sources, mostly from outside the central framework. Satellite accounts offer opportunities to clarify the problems and present them in an integrated manner. Using the rules and concepts of the SNA, a satellite account provides by definition some level of comparability, but the problems related with cross-country comparability are not solved. The contents of the health care systems to be compared at any level of aggregation are too different.

SNA 1993 presents two possibilities for the creation of satellite accounts, one close to the central framework and one with complete freedom in concepts, classifications, definitions etc. It is obvious that although the latter type opens a vast area of opportunities for creating exactly what one wants, it is not the best solution for cross-country comparability. A satellite account developed in close connection to the central framework will be more comparable.

Satellite accounting, being included in the SNA, has matured to a certain level, which entails that the chances of more uniform definitions, concepts and classifications being used have increased. Theoretically speaking, deviations in the data used for satellite accounts as well as deviations from the central framework are known and visible.

Moreover, satellite accounting by definition offers more details on the topics described than the central framework. More details in themselves are no advantage for the results of cross-country comparison, but they do allow more specific comparisons; they bring all kinds of comparisons down to a more specific level. In addition, more specification in the direction of policy and management-relevant areas of social concern offers opportunities for providing information for national goals as well as for comparisons with other countries.

However, some peculiarities can be mentioned:

- since the inclusion of chapter XXI in SNA 1993, hardly any literature has been presented on the topic of satellite accounts, more precisely on the topic of satellite accounts on health care;
- no effective introduction of the concepts and methods described has actually been carried out. Only in France a firm system of satellite accounts is established, incorporating a satellite account on health. In very few other countries has this methodology as described been put into practice. Foulon mentions that the Netherlands is the only other country that creates satellite accounts (Foulon 1982, page 59). The statistics he refers to (costs and financing of health care), however, are in my opinion not a satellite account, but a set of functional accounts on health care.

In Canada a start was made on a framework for satellite accounting. Although it was already clear from the beginning of the project that the creation of a "... fully integrated system would be immensely complex and would require extensive time and resources, ... the magnitude and complexity, ... lack of resources, ... and absence of support ... contributed to the cessation of the project" (Sunga 1986, page 278).

In Denmark (see Hansen, 1996) satellite accounts on health care were created in the 1990's as an experiment, detailing national accounts data to the level of diagnostic categories of individual hospitals. This original work can be treated as satellite accounting in SNA 1993. International comparability, however, depends on the possibilities that diagnostic related groups (DRG's) offer as a solid basis for comparisons;

- internationally, only satellite accounts on the environment are produced seriously and used on a large scale, set up by a large number of countries. Eurostat and the OECD are presently developing satellite accounts on tourism. Within the health care area the possibilities for satellite accounting seem to be winning less favour. Functional accounting is much more in the picture. This may be because the (forced) linkage to SNA is too much of a restriction.

Some specific problems still have to be solved in the general accounting framework on health care, mainly in the area of terminology and boundaries of activities of the providers:

- the first problem is in the terminology used. As part of the national accounts satellite accounts are not concerned with health but with health care. But health care itself is still not clearly described. Foulon mentioned that “SNA interpretations differ too much from one country to another, largely depending on statistical traditions and on health systems with various historical backgrounds.” (Foulon 1982, page 50). Two decades later, SNA interpretations still differ, but that is true for the whole statistical area to be described in the SNA and not specific to health care. And health care systems are still subject to past developments and “... activities of hospitals or other health institutions [vary] from one country to another.” (Foulon, 1982, page 49);
- differences in health care systems are real differences, as are differences in the totality of activities performed by the various providers of care. What is important in cross-country comparisons, is that the totality of activities to be compared has the same contents. The boundaries of the health care system have to be drawn based on the same sets of activities and the same notions of the area to be described. In this area in particular, some progress has been made in the last decade.

Within the construction of satellite accounts on health care, international agreement has to be reached on the following topics:

- first of all the boundary of the area to be included. International agreement must be reached on the inclusion or exclusion of education and research activities in the core of the satellite account, as well as on which goods and services are characteristic and which are connected. Another boundary problem is created by the supplementing of public administration with private health care administration. If no agreement can be reached, the minimum requirement is that these areas of dispute are separately visible in the accounts;
- secondly decisions are needed on which secondary and ancillary services are to be traced in the central framework and included in the core satellite;
- thirdly it is imperative to reach agreement on the boundary of health care, meaning that agreement is needed on which additional activities are to be included e.g. from household production in health for own use;
- within the branch of health and social care (NACE 85) the secondary activities outside health care are included, but the division between health and social care

activities needs to be clear. Social care activities as a secondary activity of health care providers are treated as a 'pollution' of the main activity.

In conclusion it may be stated that although satellite accounts do offer possibilities to describe in detail an area of (social) concern, the linking with the central system of SNA still may be too rigid. There appears to be a difference between the economic possibilities offered by satellite accounts and the data needed for policy solutions, which may be an additional reason why satellite accounts (on health care) are not widely constructed and used.

4.7 Final remarks

In answer to the questions presented at the beginning of this chapter, we can conclude from the relevant literature that the international community agrees to a certain extent on the contents of satellite accounts. This is not to say that there are no differences in the description of the contents and the goals of satellite accounting, but there is some agreement on the basic goals and proceedings. Another point in favour of international agreement is found in the inclusion of a chapter (chapter XXI) in SNA 1993. Opposite this notion of international agreement, it can be mentioned that SNA offers two possibilities instead of one internationally agreed solution to the problem.

The basic starting points in the construction of satellite accounts concern added flexibility to the system and the limitation of the area of study to matters of social concern. A seamless link to the SNA is a criterion for a satellite account. This does not mean the data have to be identical, but that at a certain level of aggregation the data of the satellite account must be traceable in the national accounts and vice versa. This implies that the data related to all secondary and ancillary activities have to be included in these aggregates. Increased flexibility is one of the original motives for the initial idea of satellite accounts, and is generated by the possibility of linking non-monetary data, and the possible use of additional concepts, classifications and definitions; even the possibility to add monetary data not present in the core framework is not prohibited.

Some of the shortcomings in the central system are diminished, like the rigidity of the system and the overburdening of the system with the inclusion of more details needed for satellite analysis. Another advantage is that even deviating concepts do not pose a problem for the core system if used in satellite analysis. For comparability reasons, however, deviating concepts used in satellite accounting require international agreement.

Setting up satellite accounts for areas of social concern can lead to better usability in a national perspective and thus to a demand for more and better internationally

comparable data on the same topic, which in turn can lead to better satellite accounting.

For statistical services the construction of satellite accounts can be positive because they constitute an analytical tool that can contribute to the elaboration of administrative decisions on items of social concern, while at the same time enabling analysis in the larger institutional framework of the SNA.

In practice, however, the solutions offered do not seem to be attractive for policy managers or for statistical services, because the information supplied at the macro level is not the information needed at the meso or micro level. The interpretation and application of the principles of satellite accounting could have created some doubts, which might explain why they have not been set up on a large scale. Without political support and the mutual interest of government policy departments and statistical national accounts departments, satellite accounts are not likely to be produced.

If the guidelines and starting points are followed, it will be possible to compare results at the branch level, for example health care. The inherent problems of classifying units of production is hampering comparability in the same way in satellite accounts and in the central framework. Because the location of units of production is not only determined by the activities performed but also by the local national background and history, a cluster of health care activities performed in a separate unit in one country may be hidden in a larger production unit in another and thus no longer visible or traceable.

In conclusion it can be stated that satellite accounts offer more flexibility than the central framework, and that they stress the need and offer the opportunities to expand the central framework without disrupting and overburdening it. Satellite accounts provide more insight into the problems of comparability because they contain more detail, but they do not supply a definite solution to cross-country comparability of the health care area. They cannot solve the basic problem of comparability of the contents (expressed in activity patterns) of providers, which is caused by existing differences in activities performed.

My personal opinion on satellite accounts is that although all authors agree on the principles along which satellite accounts should be produced (e.g. the linkage to the aggregates in national accounts), all these starting points are less stringently applied than necessary.

Satellite accounts as described in SNA 1993 are not very rigorous regarding the nature of the links to the central accounts, nor very specific in respect of non-monetary information to be included, nor very rigorous on the level of freedom allowed.

There are articles and educated proposals (including chapter XXI of SNA 1993), but to date no consensus has been reached on what satellite accounts are or on the minimum set of rules they should embrace. There appear to be pioneers aiming at

such consensus and rules, who at the same time present their failure to do so for many of the important questions they seek to answer.

Notes:

- 1) Production value is the value of the goods and services produced in market prices (basic prices or production cost) seen from a supply or supplier's point of view (derived from ESA 95, see also Chapter 3, footnote 11).
- 2) Market prices are ESA's basic reference for valuation. Supplies of products are valued at basic prices. Uses of products are valued at purchaser's prices in ESA 95 (see ESA 95, page 37). SNA 93 still allows producer's prices to be used in the valuation of supply.
- 3) In France a slightly modified version of the national accounts as published in the SNA is used: the enlarged system of national accounts (Le système élargi de comptabilité nationale). The use of this enlarged system may be one of the main factors why satellite accounts are, at least in part, a success in France.
- 4) OECD, Sitzung einer Ad-hoc Arbeitsgruppe über Tourismusstatistiken, Interne Arbeitsdokumente von Österreich und Frankreich, 4/5 Februar 1985.
- 5) See World Development Report 1993: Investing in Health, Worldbank, Washington, page 55, in which these items are specifically mentioned as reasons for government action in the health care area.
- 6) Expenditure (derived from ESA 95, page 51) is a way to express the exchange of goods and services, i.e. when the purchaser incurs a liability from the seller, for cash from the paying or purchaser's point of view.
- 7) The difference between these two terms is that production value is a term used in supply and hence measured in basic prices, while expenditure is a term used in uses (of the products) and hence measured in purchaser's prices. The purchaser's price (see ESA 95, page 37) is the price the purchaser actually pays for the products; including any production taxes less production subsidies (but including any transport charges paid separately).

5. *Functional classifications*

5.0 *Questions*

Can functional classifications provide a solution for international comparability of health care expenditures?

Is it possible to create a set of expenditure tables based on these functional classifications, which are comparable across countries?

5.1 *Introduction*

The previous two chapters have shown that neither national accounts nor satellite accounts can provide an adequate solution for the problem of cross-country comparability. Perhaps functional classifications¹⁾ will afford more possibilities for comparability.

Unlike the systems described in the last two chapters, functional classifications cannot be seen as a framework, but more as instruments that can be used within or outside the framework of SNA/ESA (System of National Accounts/European System of Accounts). In some countries functional classifications are treated as an integral part of SNA, while in others they are treated as supplementary to the system. Functional classifications are instruments to classify transactions. Classifications are the basis for an accounting approach; they do not stand on their own. In this chapter the possibilities are investigated of functional classifications as mentioned in the SNA, as they can also be applied outside the rigid framework of SNA. Functional classifications are considered on their own merits and as instruments in the preparation of expenditure²⁾ accounts in the field of health care, although it should be noted that while they solve the problem of identification (location of items), they do not solve the problem of valuation.

5.2 *Description of functional classifications*

There is no clear definition or description of a functional classification. In all the relevant literature (see References) the term functional is used "... because these classifications identify 'functions' in the sense of 'objectives' or 'purposes'." (SNA 1993, page 415). Additional information on the term 'functional' can be found in the Classification of expenditure according to purpose (UN 2000a). "The three words

used (function, object and purpose) convey the same meaning, [that is] the socio-economic objectives that institutional units aim to achieve through various kinds of outlays." (UN2000a, page 10). The classifications are designed to order transactions that result in the payment of money.

In the same publication (UN 2000a) the use of the classifications is linked to institutional units ³⁾, which according to ESA 1995 are economic entities that are capable of owning goods and assets, of incurring liabilities and of engaging in economic activities and transactions with other units in their own right (ESA 1995, page 8). In SNA the term enterprise is used. The term functional classification is not limited to institutional units. Functional classifications can be used to classify transactions of any distinct unit; even units that are only part of institutional units. These (parts of) institutional units serve a role in the process of economic production and consumption, in which transactions take place resulting in actual payments (expenditures).

In this study functional classifications refer to a classification of expenditures grouped according to purpose of production and/or consumption.

5.3 *The history of functional classifications*

The functional classification of inputs was introduced in the 1968 edition of the SNA (SNA 1968, page 15), in which two chapters included tables containing the classifications by purpose. In chapter 5 the classifications on the purposes of government and the purposes of non-profit institutions serving households (NPISH) are presented in table 5.3 (page 87–89, SNA 1968) and table 5.4 (page 89, SNA 1968) respectively. In chapter 6 the classification of goods and services for households by purpose is presented in table 6.1 (page 105–108, SNA 1968). The classification of functions of government (COFOG) was profoundly revised in the early 1980's, i.e. the structure was brought into line with the one used in the classification of individual consumption according to purpose (COICOP). A first draft version of the classification of outlays of producers by purpose (COPP) was presented in 1975.

At the international level these classifications are still young. For COFOG it can be noted that "... basic data on general government expenditure are normally derived from the accounts of government agencies, whose responsibilities do not necessarily coincide with any functional breakdown like COFOG." (UN 1986, page 16). Statistical offices are forced to convert available data to the COFOG classification, which is often a very laborious task, if the level of detail of the data allow such a conversion at all. For other levels of government and other government agencies the adaptation of the accounts and the data can be even more difficult.

The department of national accounts of Statistics Netherlands, supplying information based on COFOG, notes in its 1999 and in its 2000 edition (NR 1999, page 166, NR 2000, page 166) that the information presented is still experimental. In other countries, too, the publication of tables according to purpose was not yet standard practice in 2001. Eurostat collects data according to ESA 1995 based on a regulation (EU Regulation no. 2223/96, June 25 1996), under which data by COFOG must be supplied from 2002 for 1995–2000. Under the same regulation, data on COICOP must be supplied from the end of 2000 for 1995–1999 (see Annex B, Regulation no. 2223/96).

SNA 1993 devotes a chapter (chapter XVIII) to functional classifications. Because these classifications were not completely satisfactory, the ISWGNA (Inter-Secretariat Working Group on National Accounts) already asked for a revision of them in 1995. The Organisation for Economic Co-operation and Development, OECD (in close co-operation with Eurostat) and the UN Statistical Division took it upon themselves to do this. In 1999 the revision of the classifications was approved by the UN (United Nations) Statistical Division. In 2000 a publication was presented on the four functional classifications (COFOG, COICOP, Classification of Purposes of Non-Profit Institutions Serving Households, COPNI and COPP) used in the area of national accounts.

The term functional was introduced in SNA 1968; the 1993 version of the SNA treats the terms functions, purposes and objectives as synonyms. In the 2000 publication the term function is consistently replaced by the term purpose.

As already mentioned in chapter 3 (see table 3.1) of the present study, the ESA does not explicitly mention the classification of purposes of NPISH nor the classification of purposes of producers. The other two classifications (on functions of government and consumption of households) are included in the European framework of statistics. ESA 1995 does not contain a separate chapter on functional classifications. COFOG and COICOP as tools to be used are mentioned in Section 9.51 on page 223.

5.4 Objectives of functional classifications

Functional classifications (inside or outside SNA or ESA) are an instrument for classifying transactions of producing units or consumption of households. These classifications are an inherent part of the SNA, with all its rules and definitions. From a SNA point of view, functional classifications can be used for the description of specific transactions of producers (in COPP). Functional classifications can also serve to classify and describe the specific transactions of three sectors distinguished in the national accounts framework, i.e. government, non-profit institutions serving households (NPISH) and households (consumers).

Functional classifications serve three distinct objectives, all related to expenditures:

1. to separate collective services from individual (or individualisable) goods and services. Individual goods and services are deducted from government consumption expenditure and added to individual consumption expenditure (see SNA 1993, page 415, par. 18.2);
2. to provide statistics considered to be of general interest (see SNA 1993, par. 18.3): national welfare statistics (e.g. food, health and education services), or statistical information on so-called merit goods (e.g. again health or education), which are supplied by government below cost price to encourage their use. Another example of general interest is information on outsourcing of business services which were mostly carried out as ancillary activities (e.g. buying in catering, cleaning or transport services);
3. to enable users "... to recast key aggregates [...] for particular kinds of analyses." (SNA 1993, page 415, par. 18.4). Examples of possible analyses are the creation of human capital based on education expenditures to be used in labour productivity studies, or the use of intermediate consumption on R&D expenditures as capital formation in economic growth analysis (for more examples see UN 2000, page 11).

The functional classifications can also be used outside the rigid framework of SNA/ESA. They are instruments to order transactions of individual units of production or consumption. For every unit of production – whether it falls under government, non profit institution or enterprise – transactions, including transactions in the area of health care, can be separated using the functional classifications. The same is true for the consumption transactions of households: the functional classifications can distinguish these transactions by the purpose they serve.

Although all three above-mentioned objectives for the classifications may be relevant in the present study, for comparability the objectives for which the classifications are used within a country are not important. For this study a fourth objective should be identified: comparability across countries. For comparability it is important that the results based on the classifications are in line with each other.

5.5 *Functional classifications on health care*

SNA 1993 and the UN publication on classification by purpose (UN 2000a) distinguish four classifications (see also table 5.1), three dealing for final consumption and one for intermediate consumption (or production) expenditures. All four functional classifications are concerned with expenditures, be they consumption expenditures or expenditures of producing units for specific functions. In theory, looking at production and consumption, government and NPISH could be divided according to the producer outlay classification. Because

almost all production of government and NPISH is considered to be consumption expenditures, it is more practical not to divide these two producers according to a producers classification, but according to a consumption classification. The majority of the production in ISIC 85/NACE N – health and social care – concerns the production of services, which cannot, by definition, be produced for stock. Basically for the results no difference will be noticeable.

In all four functional classifications health is listed as a function or purpose to be distinguished separately. Although none of the classifications provides a definition or even a description of health, all four supply a list of institutions and/or goods and services supplied by providers belong to the category health care. A unit classified as belonging to a certain production group does not mean that the totality of activities performed by that unit is the same in all countries. Thus goods and services determined as belonging to the health care field are not automatically the same internationally (see chapter 11 on the Common Comparable Package).

In SNA 1993 the structures of the classifications are outlined at one-digit (COPNI) or two-digit level (COFOG, COICOP and COPP). In the UN publication (UN 2000a), in collaboration with OECD and Eurostat, all classifications are defined at the three-digit level ⁴⁾.

Table 5.1
Functional classifications on health care

Producers/Expenditures	Market/Non-market		Government		NPISH ¹⁾
	ISIC 85.1	Other ISIC	Collective	Individual	
Producers COPP ³⁾	X ²⁾	X	[X]		[X]
Consumption COFOG ⁴⁾					
– collective			X		
– individual				X	
COPNI ⁵⁾					X
COICOP ⁶⁾	X ²⁾			X	X
Consumption by households	X	X	X	X	X

¹⁾ NPISH: Non-Profit Institutions Serving Households.

²⁾ Production of units in ISIC 85.1 Health care is not identical to consumption as registered in COICOP under health care.

³⁾ COPP: Classification of Outlays of Producers by Purpose.

⁴⁾ COFOG: Classification of Functions of Government.

⁵⁾ COPNI: Classification of Purposes of NPISH.

⁶⁾ COICOP: Classification of Individual Consumption by Purpose.

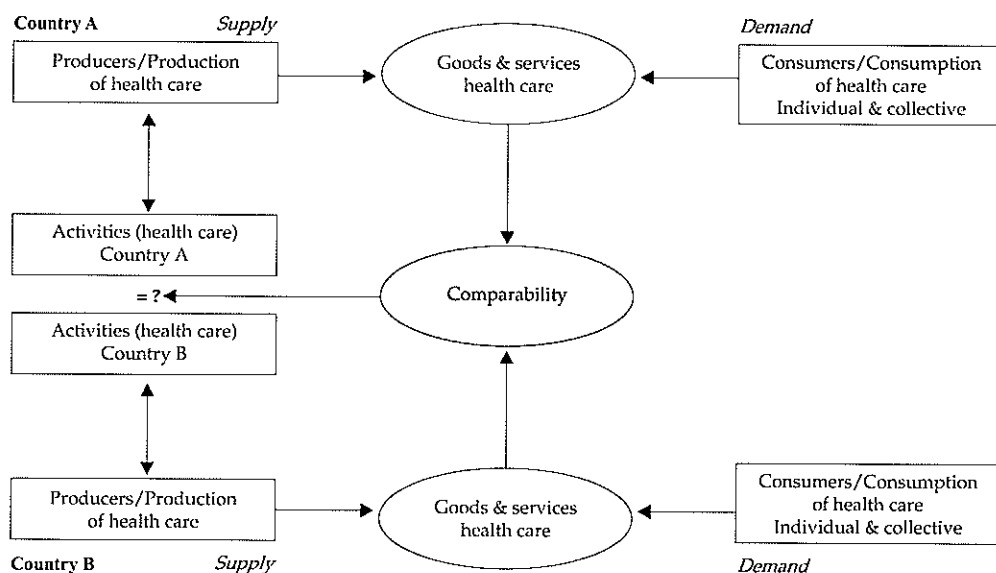
5.6 Functional classifications on health care and comparability

All four functional classifications contain the function health and are concerned with expenditures, final consumption expenditures or intermediate consumption expenditures. As argued in chapters 1 and 2, the term health should be replaced by the term health care. The expenditures are related to health care provisions, services or activities. Indeed it is hardly possible to spend money on health; money is usually spent on health care.

In theory it is possible to link all four classifications. Production of the purpose health care totalled over the three producing classifications (COFOG, COPNI and COPP), corrected for the intermediate health care consumption of own activities and for collective health care consumption in COFOG, should equal the consumption of the function health care as presented by the consumption classification (COICOP).

This is true for every country. For the purpose of international comparison the production side of this equation needs more attention (see Figure 5.1.). The goods and services produced and supplied are not the ultimate limiting factor for comparability. Goods and services are produced by production units inside and outside the specific ISIC/NACE classes of health care. For cross-country comparability it is necessary that packages of goods and services – and for that reason packages of activities performed by the producers – are the same.

5.1 Comparability and functional classifications



Based on this theoretical equivalence and the notion of health care, it might be concluded that for comparable health care data, the functional classifications can provide a contribution to the solution of the problem.

5.6.1 *Advantages or favourable conditions*

There are a number of arguments in support of this statement:

- all four classifications use transactions as the basic building blocks. COFOG and COPNI use transactions to classify expenditures, among which consumption expenditures on the function of health care. COICOP and COPP use outlays on individual consumption, among which health care (in COICOP) and production of health care as an ancillary activity (in COPP);
- SNA 1993 introduced the new concept of actual final consumption. “The immediate consequence of the introduction of that concept ... was that for all expenditures related to social transfers in kind ... COFOG, COICOP and COPNI had to be harmonised.” (UN 2000a, page 29). Social transfers in kind are treated as individual consumption and are therefore separated from collective consumption expenditures. A transfer of these individual consumption items from COFOG and inclusion in COICOP will thus be possible. All expenditures related to NPISH are by definition individual, so these expenditures can be transferred to COICOP as well. By definition harmonised classifications – unanimously and identically applied – are better for comparability than non-harmonised classifications, because they are consistent with each other and thus have already achieved (a certain level of) comparability;
- for cross-country comparability, one advantage is the treatment of the outlays in COICOP. Outlays that possibly serve more than one function, the so-called multi-purpose goods and services, need to be classified according to the division that represents the predominant purpose. If the predominant use varies across countries the items are “... assigned to the division that represents the main purpose in the countries where the item concerned is particularly important.” (UN 2000a, page 20). Comparability (in SNA terminology) across countries is improved because the functions or purposes of the so-called multi-purpose goods are harmonised and limited to only one function or purpose. For national purposes, however, the reality – both economic and social reality – concerning expenditure of these functions, could be missed by this ‘majority’ rule.
An example of a multi-purpose good is a bicycle, which can serve the purposes means of transport, means of leisure activity, sports item, toy or means of health enhancement. The international agreement to classify a bicycle as a means of transport, creates better cross country-comparability. As a consequence it reflects less well the economic reality in countries where bicycles are not primarily used as a means of transport;
- a last consideration in favour of the use of these classifications might be that they can be used outside the framework of SNA to classify transactions of producers and consumers.

5.6.2 *Disadvantages or unfavourable conditions*

Opposite these favourable conditions for comparability, however, there are some disadvantages:

- government subsidies are not treated consistently with regard to comparability. In general, subsidies should be classified under the function they serve. Programmes, subsidies, loans or grants concerning labour and the labour market, however, are not classified by the purposes they serve, but under general labour affairs. These financial flows connected to the labour market are an exception to the general rule, because they do not focus on any single industry, but are meant to serve all industries or sectors in the economy. Completeness of the expenditures on the function, including the function of health care, is not guaranteed. Cross-country comparability is certainly influenced, because the numbers of employment-related subsidies and the division across sectors or industries differ between countries and both factors influence the outlays by purpose that are to be compared;
- a bigger disadvantage in relation to comparability are the classification procedures. Although basically transactions need to be classified, it is not always possible to use transactions. If they cannot be used, the rule is that the producing units are the units of classification (COFOG and COPNI); and if production units provide more than one function and individual assignment of the functional codes is not possible, the majority rule is applied (see UN 2000a, page 13 and page 23). The use of units for classifying, especially if they serve more than one function and the majority rule is applied, is detrimental for comparability of the results. First, the 'purity' of the functions is blurred by secondary and ancillary activities; and secondly, the composition of pure functions and secondary activities in units of production is not the same across countries;
- the third disadvantage concerns the treatment of outlays in COICOP. Mixed purpose goods and services (consisting of more than one specific type of good or service) need to be broken down into their respective purposes as precisely as possible and "... consistent with practical considerations of data availability." (UN 2000a, page 21). The latter, in particular, is very detrimental to cross-country comparisons. Data availability on functions is not and cannot be the same across countries for all purposes and for all goods and services in a given period of time;
- one item not mentioned but crucial for international comparability is the way the data are gathered. In cases where household budget surveys are the basis for COICOP, the sample of households is representative for the population as a whole and the sections for which the population was stratified. The results of the survey are reliable within the margins set for the sample, but not for other subsets of the population if the population was not stratified according to these criteria. An additional problem is that not all goods and services as collected in a household budget survey have a unique position in the classification of purposes used in COICOP.

5.7 *Conclusions on functional classifications as an instrument for comparability*

The classifications by purpose as developed in the recent past are linked to the framework of the national accounts, which gives them a solid basis for analysis. Furthermore they offer the possibilities to limit the analysis to a very small area, for example health care expenditures. Within the boundaries of the classifications, there is freedom to handle the data.

On the other hand, the choices available in the classification rules, although offering the possibilities of completeness in the description of the area under study, are detrimental for comparisons. Choices of classification procedures within a single classification do not guarantee comparability of the results. Options for different classification rules across classifications handling the same types of information (like final consumption expenditures) make the problems concerning comparability more complex. These opportunities in classification procedures across countries do not assist cross-country comparability, but are detrimental for it, even when a small item like final consumption expenditures in health care is compared.

Although in the latest version on the functional classifications, the terminology of purpose is used, the basic distinction or even controversy between COFOG on the one hand and COICOP on the other is not eliminated. COFOG basically deals with functions or activities financed by government bodies and COICOP deals with household consumption expenditures on goods and services.

A definite conclusion on the usefulness of these classifications by purpose cannot be made yet, because they have only recently been made consistent with each other and are being implemented. However, in theory these classifications offer real possibilities to improve comparability, especially of health care expenditures on condition that for every classification one single choice is made from the opportunities offered in the classification procedures.

This latter point does not mean that the use of these classifications would solve the entire problem of cross-country comparability. Having the same units across countries or having the same totality of goods and services used for health care does not provide complete comparability. What is required is the same set of activities performed by production units.

5.8 *Final remarks*

In answers to the questions presented at the beginning of this chapter, we can state that in theory the classifications by purpose as presented do offer possibilities to

enlarge international comparability of health care expenditures, especially if these classifications are seen as instruments to ease the problem of cross-country comparability of expenditures by purpose. However, choices for individual countries in the procedures do not increase international comparability, especially not if different options are available for every classification. Freedom within the classification rules of one classification and freedom to make different choices on the classifying process for different classifications do not improve comparability of functions across countries. In my opinion the usefulness of functional classifications would improve if the international community decided to use the units of production as starting points in the classifications. In this way the link between the units of production and the activities performed by these units could become much stronger. And furthermore this would ensure that the goods and services provided by these units are comparable across countries. As a first step international agreement on the way the procedures need to be handled is necessary.

Theoretical possibilities need to be tested in practice to prove their usefulness. A second point is that more countries need to gain experience in using the functional classifications, either within the framework of SNA or within the framework of health accounting (outside SNA). These classifications need to be transferred from the experimental stage to a regular collecting stage to be useful in cross-country comparisons.

Basically, given uniform solutions for the problems and options mentioned above, it should be possible to create tables on expenditures using functional classifications which are comparable across countries. However, one condition is of eminent importance in the process. The units used in calculation by function, and the functions defined need to contain the same set of activities across countries.

Notes:

- 1) The Statistical Commission of the UN (see UN 2000a, page 9) recommended the use of the term "classifications of expenditure according to purpose", but for reasons of readability the term "functional classifications" is used in this study. Functional classifications of expenditure are not to be confused with functional classifications of units, in which statistical units are classified according to activities performed. Opposite this classification is the institutional classification, in which statistical units are classified according to their main activity.
- 2) Expenditure (derived from ESA 95, page 51) is a way to express the exchange of goods and services, i.e. when the purchaser incurs a liability from the seller, for cash from the paying or purchaser's point of view (see also notes 6 and 7 in chapter 4).

- 3) For an exact definition of the term institutional unit see ESA 1995 page 19. In SNA the term enterprise is used to describe an institutional unit in its capacity as a producer of goods and services (see SNA 1993, page 115).
- 4) In the UN 2000a publication the highest (first) digit level is presented as the two-digit level, because in some classifications more than nine items had to be listed at highest level of aggregation. Consequently in every subdivision a higher digit level results.



6. *Health accounts*

6.0 *Questions*

Is there international agreement on the construction and use of health accounts?

What are the basic criteria and starting points used in health accounting?

How do these criteria and starting points relate to comparability at branch level?

Do health accounts reduce the shortcomings in the core system of the national accounts, and if so to what extent?

6.1 *Introduction*

The previous chapters on national accounts, satellite accounts and functional classifications have shown that although each of these international frameworks offers some possibilities to improve comparability for the health care branch, the possibilities are limited and do not solve the basic problem conclusively, i.e. the differences between sets of activities performed by the various units of production in the various countries. In this chapter the usability of (national) health accounts – also called functional accounts or functional approaches – for international comparability is discussed and evaluated. Health accounts as discussed in this chapter can not be seen as being part of existing international frameworks. On the other hand these accounts have been constructed and implemented by a large number of countries using the same basic principles, which makes them part of an accepted practical framework.

As in the other chapters the terminology used here – health account – should be read as health care account, because this study focuses on the health care branch and not the wider area of health.

In order to be useful any accounting framework must relate to national accounts, in which the broader economic activity of the nation is presented. (National) health accounts are a kind of satellite account using the same basic elements of national accounting to describe health care ¹⁾. Because they are created to show the sources and the destination of health care funding ²⁾ the emphasis in health accounting is on expenditures ³⁾. Flows of money are tracked from public and from private sources (insurance and household out-of-pocket) to health care providers. Like satellite

accounts, health accounts can also incorporate other dimensions of the health care system (branch employment, production, beneficiary groups, regions, etc.). Although for some people (see Waldo ⁴⁾, page 1) health accounting is typed by using matrices to estimate and show the results, in my opinion matrices are useful in the construction of health accounts but not crucial. Matrices can be of assistance in the calculations and in the determination of the interrelationships and consistency of the material used.

Definitions and descriptions of health accounts are scarce. A good description is provided in the World Health Report 2000 (WHR): "National Health Accounts are designed to be a policy relevant, comprehensive, consistent, timely and standardized instrument that traces the levels and trends of consumption of medical goods and services (the expenditure approach), the value-added created by service and manufacturing industries producing these commodities (the production approach) and the incomes generated by this process as well as the taxes, mandatory contributions, premiums and direct payments that fund the systems (the financial approach)." (WHR 2000, page 149). Although the goal of (national) health accounts should be to create a complete system (including production, income, consumption and financing), in some countries the accounts on health care are limited to the expenditure approach. A number of elements mentioned here will be discussed in this chapter. For a more detailed discussion of national health accounts as used by World Health Organisation (WHO) see e.g. Poullier 2000 and Poullier 2002.

6.2 *History of health accounts*

An interest in and the development of health accounts in individual countries already emerged in the 1950's and 1960's, over four decades ago.

The Netherlands already performed the first studies on the level of the expenditures of health care in 1953 (CBS 1953), resulting in a yearly available, continuous and consistent set of data on health care from 1972 onwards.

France also started developing health accounts in the 1950's. The French attempts in the construction of health accounts resulted in a complete consistent system of satellite accounts on health, first published in 1979 (see chapter 4).

The USA was also one of the forerunners in the development of health accounts after World War II, work that was boosted in the 1960's with plans to make health care available to large groups in the population: the creation of the public programmes Medicaid and Medicare.

Since that same period a large number of countries have become involved in health accounting. All of the developments mentioned so far were limited to national data and national use of the constructed tables of information and therefore by definition not comparable. There are no agreed international framework, classifications, boundaries or definitions of terminology in health care accounting.

However, interest in the field of health, health care and health accounts was not limited to national governments and national agencies. International organisations also showed interest in the topic.

WHO⁵⁾ had showed already a great interest in comparable health expenditure in the 1960's. First attempts to obtain data on health care expenditure were made by B. Abel-Smith (Abel-Smith, 1963, 1967). At the same time comparability of the data collected became of prime importance. By means of structuring the questionnaires sent out to the participating countries, attempts were made to create a comparable set of starting points in the data collection by specifying in detail the data requested for various items. However, these attempts were isolated in the sense that these did not result in significant co-ordinated international efforts in the decades to follow.

In the late 1970's and early 1980's the Organisation for Economic Co-operation and Development (OECD) started creating its health database⁶⁾, prompted by the need of national agencies (mostly ministries of health) to compare the level and growth of health care expenditure with other developed countries.

Later on in the 1980's the World Bank⁷⁾ became interested in health accounting, because of increasing requests from developing countries for financial support. Such support requires a more specific and systematic compilation of health financing information.

In the European Union (EU) – not an international but a supranational organisation, having power to endorse actions decided upon – the interest for health accounting is more recent. In the treaties of the EU the responsibility for health and health care is left to individual member states. Officially, with the start of the Health Monitoring Programme at the end of the 1990's, the EU stimulated the member countries to create health accounts, and strive towards comparable results across member states.

In the 1990's more interest emerged from both developed and developing countries for health, health care and expenditure in an international context, with an eye to improving the management and control of health care as well as cost containment. Although medical science, the core of health care, has always been an international issue, it is only part of the totality of health care activities. Most actions and activities in the health care area are not internationally anchored, but have a firm local basis.

In the period described, cost containment was a serious topic in the developed world, and this required health data. In the developing countries the challenges were (indeed still are) to create health care systems that are equitable and sustainable. The use of health accounts cannot be treated independently of the control of these accounts. In general national control fosters national use just as international control fosters international use. The increased interest in health care expenditures coincided with the revision of the SNA (System of National Accounts) in 1993 in which the more flexible system of satellite accounts was introduced as part

of a solution for lacking international comparability. Satellite accounts do not always meet the information need⁸⁹ of policymakers in the health care area, so there is still a need for health accounts.

6.3 *Aims and objectives of health accounts*

Although various reasons can be mentioned for the creation of national health accounts, in principle all of them are based on a need by health care managers and policy makers for information on health care⁹⁰, both at the national level and the international level. If a system of health care needs to be constructed, if changes are proposed, if the level or growth of health expenditure needs to be controlled, information is imperative.

National health accounts are designed to answer three major questions (see e.g. Griffiths 1983, page 46) concerning:

1. the total amount of expenditure on the health care branch;
2. the sources of funding and the amount provided by each source;
3. what the money is spent on.

A fourth question, not mentioned by Griffiths, has to do with the distribution of the benefits, i.e. who receives the goods and services provided. Implicitly all authors and organisations in health care use these questions as starting points in the construction of a national health account. The interpretation of the various aspects and the sequence in which the questions are handled may differ. The lack of international agreement on concepts, definitions, nomenclatures and classifications, and terminology used is one of the most important reasons for this. This lack, already explicitly stated by Abel-Smith in 1963 (Abel-Smith 1963, page 9), is still present (Berman 1999, page 47).

Originally only national information was necessary because the problems to be solved were national by nature. Later on it was realised that developments and solutions arrived at in other countries could be useful for finding solutions for national problems. Furthermore it is easier to present and defend a decision if it is backed by experiences of other countries. The use of experiences of other countries by any national party (political, governmental, statistical or otherwise) makes it more important for the data presented to be comparable.

6.4 *Starting points and criteria*

As this chapter deals with the construction of health accounts for primarily national purposes, these accounts can be categorised as national health accounts. This

national starting point in the construction of health accounts might limit the usability in an international comparative framework, but it does not diminish the quest for comparable information by international organisations nor the mirror information by national users and authorities.

6.4.1 Criteria

For all statistics – irrespective of their subject and how the information itself is handled – and in all countries a few basic criteria in the treatment of information are taken into account to make the information usable. These criteria are part of the encompassing term quality, whose relevance for comparability is described in chapter 2 of this study. The criteria are also part of a longer list of criteria mentioned by Eurostat (UN 2000 b, page 1).

These basic criteria are:

- completeness of the information presented;
- consistency between the items presented;
- coherence between information parts and with existing systems of information;
- comparability in time;
- timeliness of presentation and dissemination of the information.

The fourth ‘c’ in the list, the ‘c’ for comparability in time is only relevant if the (preceding) criteria are met. Without completeness, coherence and consistency, comparability in time is meaningless.

Although not often explicitly mentioned, comparability in time is of prime importance for statisticians and for users of statistics. It is a difficult aspect to determine. The visibility of breaks in series, for example following the inclusion or exclusion of certain types of institutions or activities, depends not only on the relative magnitude of the change compared with the totals at the time of the break, but also on the relative importance of the real change after a meaningful time span (e.g. a few decades).

6.4.2 Features and characteristics

As stated in the first Dutch study on health accounts “... the relations between the various institutions participating in health care are so complicated, that a special ordering framework had to be applied to clearly structure this complicated network of financial flows. Statistics Netherlands used for that reason the System of National Accounts, which is used to summarise all transactions between all relevant sectors of the economy.” (CBS 1953, page 6). The basic sources of information available, however, did not allow for a complete implementation of SNA principles in the national health accounts. Even if the complete implementation of SNA principles had been one of the goals of the Dutch health accounts (which it was not), it is questionable whether the results would have been better, both from a national viewpoint and from an international comparable one.

In the USA national health accounts were developed to be able to control and manage the extensive public health programmes Medicare and Medicaid. The goal of health accounting there, however, has never been to approve or disapprove health care expenditure. The first USA health account was constructed by national accountants, of course using the principles and starting points of the SNA.

Other countries had similar reasons for using the SNA as a basis to construct health accounts without completely implementing it. National accounts and health accounts have many concepts and methods in common. They are two separate but parallel approaches. They differ because they were developed separately and for different purposes. National accounts were primarily set up to present relations between one branch and the national macro-economy. A health account describes the flows of inputs and expenditure between various institutional elements of the health care system. Health accounts stress expenditure instead of the valuation of economic activity, which is a goal of national accounts. In national accounts the main activity of production units is a leading principle, while in health accounts all nationally defined activities in the area of health care are the core principle. National accounts clearly distinguish between production and expenditures, which are mixed in health accounting (probably because most managers and policy makers look upon the health care branch as a spending mechanism instead of an economic force).

The strength of national accounts lies in their facilitation of economic analysis and cross-national comparability of national macro-aggregates. SNA is mainly used for analysis across sectors, not within a sector. The use of the central framework, its definitions, classifications and terminology is often in conflict with the needs expressed by health managers and thus limits the relevance of results based on national accounts. Circumventing the disadvantages of the national accounts for health managers was what prompted the development of national health accounts. The same reasons, however, are possible disadvantages in international comparisons, because national uses deviating from national accounts practices, can create cross-border incomparability.

In every country setting up health accounts, the same goals and purposes are set, but the criteria used to define, for example, the boundaries of health care, methods, data, calculation and estimation methods can differ to some extent, because no agreed international framework or set of principles is available or used in constructing the health accounts.

6.5 *Health accounts framework: results*

The development of health accounts with international comparability possibilities can be traced back to two studies performed by Abel-Smith for WHO in the 1960's. In respect of the data presented by the participating countries in these studies, he concluded that "... too much weight should not be placed on some of the figures in view of the difficulties that some countries had in distinguishing teaching and

research from other health services and in separating public health from other services. Moreover some countries were unable to make an accurate division between hospital and non-hospital services" (Abel-Smith 1967, page 55). He had mentioned the same points in the pilot study (Abel-Smith 1963, pages 50–53).

Abel-Smith tried to achieve comparability across countries for the first time using the principles of national accounts applied to the field of health care by using a single questionnaire as the instrument to collect the information needed.

In spite of its being an agreed framework, SNA created some of the problems in this study which hampered comparability. Disadvantages are that SNA does not clearly specify what health expenditure is, and it is not detailed enough in its breakdowns. Another disadvantage is that the flow of money in SNA is not completely identified in the detail and the relations between the principal categories in health care as requested for health care analysis (see Abel-Smith 1967, page 100).

Abel-Smith's method was "... to analyse and adjust existing data ..." (Abel-Smith 1963, page 10). Adjusting existing data is a well-known method to create cross-country comparability, and is used by almost all international organisations. Adjusting existing data by a researcher may involve moving data from one sub-specification to another to create more homogeneity in these sub-specifications. Adjusting can also mean changing the data supplied just to get 'better' results across countries.

In my view, Abel-Smith's approach was the first practical attempt to create more comparability in data on health care. However, basically the results he showed were in essence incomparable.

Many of the initial health accounts estimates available were developed by consultants before specifications and classifications of the type proposed by OECD (see chapters 7 and 12) were readily available. Few of these accounts were directly linked to policy needs, or properly disseminated, resulting in a sub-optimal divide between production and use of the results.

Three decades later at the end of the century the results in terms of comparable data are still poor. "Not surprisingly, the value of these individual studies has been modest, [because] methods vary." (Berman 1999, page 50).

We have to keep in mind that the huge changes required by the implementation of SNA 1993 started nearly a decade ago, and in half of the world the full implementation is not yet in sight. National health accounts, building on the principles of the new SNA, will have to wait a few years before these principles are fully implemented.

6.6 Health accounts and comparability

As health accounts are basically constructed for national purposes, not for purposes of international comparisons, usability in international settings is not the main

objective. Cross-country comparability is almost never mentioned in national statistical criteria for information. Breaks in series mentioned in the previous section influence cross-country comparability as well, both positively and negatively. Breaks in national series may be prompted by international guidelines or classifications resulting in more comparable results across countries. On the other hand they may result from national political decisions contradicting internationally agreed definitions or guidelines, and thus resulting in less comparable data.

OECD (OECD 1993, page 43) mentions comparability as a criterion for internationally used data in 1993, the first time that comparability as a criterion is explicitly formulated. It will be clear that for international organisations comparability is a fundamental criterion: they could not do without cross-country comparability.

Cross-country comparability is important for most international organisations. Decisions of international bodies to change systems (e.g. SNA) can create a conflict of interest between cross-country comparability and comparability in time. Comparability in time and across countries can be hampered by national decisions to change the systems for which the information is gathered, as well by decisions of international organisations on internationally agreed classifications and systems. In the transient period, after the decisions have been taken and put into force and before all the countries have implemented the new system, cross-country comparability will be less than before.

A paper submitted by Eurostat for the Conference of European Statisticians in 2000 (UN 2000b, page 1) contains a list of "... desirable attributes of official statistics ..." or criteria for statistical results¹⁰. The most important of these attributes or criteria are timeliness, accuracy and comparability. Although all the attributes mentioned in the paper may be in conflict with each other, they can be traded against costs, but also against each other. The purpose for which the data are needed determines the relative priority of the criteria.

In the 1960's there were few national studies in the area of health care expenditure; today almost every country monitors health expenditure on a more or less regular basis. But just as in the 1960's "... the information published in one country cannot be reliably compared with that published in another, as the nomenclatures used to describe health services varies from one country to another." (Abel-Smith 1963, page 9).

A next point already mentioned by Abel-Smith concerns the definitions and boundaries used in the health services. "The field of health services is ... a continuous spectrum, and it is extremely difficult not only to define the field of health services but also to break it down into clearly defined categories." (Abel-Smith, 1963, page 24). One might say that the basic problem of comparability caused by the

differences in activities performed by units of production in the health care field across countries is already presented here.

Some developments since the 1960's have clearly provided better and clearer descriptions and definitions of health care, but in international comparable studies still "... the need for concepts which were in themselves meaningful [need to be] balanced against the need for definitions which [have] some reasonable prospect of being applied to existing data." (Abel-Smith 1963, page 24).

In itself, though, application of concepts in existing data structures is not enough. Concepts differ between countries, as do the accounts which are by nature by-products of the administrative processes. "To design one questionnaire [to be used in the participating countries for data collection] ... involved major compromises between conflicting interests." (Abel-Smith 1967, page 29). "The information which is sought [in international comparisons] will not emerge ready-made, ... adaptation and adjustment of existing data is inevitably required." (Abel-Smith 1963, page 24).

However, "... international comparisons can be misleading, because of the difficulties mentioned above of defining health expenditure on a comparable basis." (Mach & Abel-Smith 1983, page 24). If on the other hand "... such [health care] services had been uniform throughout the world, a study of this kind would hardly have been necessary." (Abel-Smith, 1967, page 29). By 2001 not much had changed. In both studies Abel-Smith concludes that "... the reported deficiencies in the returns [of the questionnaires] could not have produced any major lack of comparability in the principal figures." (Abel-Smith 1967, page 36). So, despite all the differences in definitions and concepts, estimating and calculating methods as well as boundaries and contents of the health care systems, the data presented are 'comparable'. However, in chapter 2 on sources and methods he already states that "When there appeared to be a serious lack of comparability in the figures provided by a particular country, they have been omitted from the appropriate tables, or attention has been drawn to the lack of comparability." (Abel-Smith 1967, page 36).

Berman distinguishes three lines of research in health accounting emerging from this first attempt by Abel-Smith to achieve international comparability of health care expenditure (Berman 1999, pages 48, 49). The first one, and most important for the Western, developed world, is the line followed by the OECD in the creation of the health database. This database and its use for international comparability will be discussed and presented in the next chapter. The second is USA's national health accounting methodology (the Harvard NHA methodology), which presents the information on the flow of funds in a matrix form and which is especially applicable in developing countries. The third line of health accounting followed national accounts and led to the construction of satellite accounts.

Berman forgets that WHO's first attempt originated from the experiences obtained in various countries with the development of national health accounts. Health

accounts already existed before Abel-Smith conducted his research; they are not just a consequence of it.

Basically until the 1980's health accounts were set up as so-called T-accounts¹¹⁾, which are still used in traditional bookkeeping and in SNA. "The systematic compilation of sources and uses is not specially recommended." (Berman 1996, page 11). So there is no inherent force to ensure that all the sources are included in the health accounts.

The matrix approach which is basic to the USA's NHA (National Health Accounts) methodology has two major strengths. "First it imposes a rigor on the analysis and second it needs to include the entire range of spending through which the patterns and relationships become more clear." (Waldo 1999, page 1). "The weakness of [the Harvard method on] NHA remains its lack of international comparability and internal consistency." (Rannan-Eliya 1997, Abstract).

The national health accounts created before the Harvard methodology, although different in methodology, did not cease to exist but evolved and converged in time. What is more, the type of health accounting approach taken does not solve the problems of incomparability. Just to mention a few problems all methods encounter when creating cross-country comparisons: boundary problems, disaggregation (in subcategories) and comparison, definitions, and calculation and estimation methodologies used.

6.7 *Final remarks*

When Abel-Smith started comparing the expenditure of health care systems of different countries, there was no agreement at all on definitions, boundaries, concepts, reporting systems, calculation methods or any other important item in this area. Without going into the new developments starting in the last decade of the previous century, some progress has been made. It is commonly accepted that health accounts need at least to answer the flow of funds questions (see Section 6.3.). There is, however, no agreement on the interpretation and handling of these starting points. And basically there is no need to agree, because national health accounts are primarily created to serve national goals.

Although no agreement exists on the methodology to be used in national health accounts to achieve more cross-country comparability, all experts agree that SNA has its limitations when used for health accounting. SNA is based on the principal activity classifications, designed to present the relations between the sectors of the economy. SNA prescribes specific treatment of nominal charges on health services, being transfers to government and not direct payments. Also the treatment of the informal sector and the relations between the household sector, the government sector and the non-profit institutions in the SNA limit its usefulness for health

accounting. Health accounts serve a purpose different from the more general economic ones on which SNA is based.

Health accounting methods do not lead to a decline in the shortcomings of SNA; the shortcomings of SNA are circumvented by using specifically created methods for solving health care problems in a statistical sense.

When performed properly, health accounting uses the same set of starting points and criteria as any systematic description of economic phenomena. The interpretation of these starting points and criteria and the use of the calculation and estimation methods can vary across countries, resulting in deviating results.

Cross-country comparability will not be achieved by using health accounting methodologies, unless these methodologies are treated in a special way, and changed just to produce comparability. Changing existing data sets is extremely difficult, time consuming and requiring a great level of co-operation between various organisations and countries. Because of possible policy implications, it is also politically sensitive. Health accounts created to solve national problems will not solve the problem of international comparability of health care data.

Creating an internationally agreed system of accounts to be used for cross-country comparability is a much better solution. The OECD created a System of Health Accounts (for a discussion see chapter 12) in close co-operation with Eurostat, the statistical office of the European Union (EU), to be implemented in all EU member states. Internationally agreed systems of health accounting are very useful, especially when cross-country comparability has to be maintained on a regular basis.

Notes:

- 1) Satellite accounts are distinct because the totals of these accounts are inextricably linked to national accounts.
- 2) A distinction is necessary between directly financing units and the penultimate entities providing resources to the system.
- 3) Expenditure (derived from ESA 95, page 51) is a way to express the exchange of goods and services, i.e. when the purchaser incurs a liability from the seller, for cash from the paying or purchaser's point of view (see also notes 1 on production value and 7 on valuation in Chapter 4 and Chapter 3, note 11 on valuation).
- 4) Waldo, D.R., *Creating Health Accounts for Developed and Developing Countries*, Worldbank Human Development Department, 1996.
- 5) Monetary variables like expenditure are not the prime objective of WHO, these variables are used to support health goals, which does not mean that WHO is not engaged in providing monetary data related to health.

- 6) OECD, *Measuring Health Care 1960–1983*, OECD, Paris, 1985.
- 7) *World Development Report: Investing in health*, Worldbank, Washington, 1993.
- 8) A central term used in Satellite Accounts, being part of SNA, is production value (production valued at basic prices), a term that not necessarily coincides with expenditures on health care (a term used for uses of products valued at purchaser's prices), in which managers are interested in, even when the boundaries are the same (see also notes 1, 6, and 7 in Chapter 4).
- 9) See for a description of the relation between health accounting and policy uses an article by Bonte, J and P. van Son, *Health statistics in the Netherlands for Health Policy and Evaluation: Experience and Developments*, *Statistical Journal of the United Nations ECE* 5 (1988), pages 123–133.
- 10) The complete list of attributes presented contains: accuracy, rapidity, timeliness, comparability in time, comparability among countries, detail, consistency, coherence, reliability, adequacy, relevance, transparency, handiness, or openness (UN 2000b, page 1).
- 11) 'T'-account is a tool in which incoming and outgoing items are separately presented leaving a balancing item on either side of the vertical line of the 'T'.

7. *OECD health database*

7.0 *Questions*

Which criteria and starting points were used in the construction of the OECD health database?

Are these criteria still the same in 2002 as when the database was constructed in the early 1980's?

Do and did countries supply the information on expenditure requested in the way the templates described them, and if not are the deviations important in relation to comparability?

Can the problems resulting from eventually deviating reporting behaviour be redressed to improve comparability?

Is comparability as presented in the OECD health database considered sufficient in relation to the goals set for the database? Is cross-country comparability one of the goals of the database?

7.1 *Introduction*

The last chapter on existing frameworks in the context of international comparability of health care expenditure deals with the OECD health database. This is the logical next step after the previous chapters as some countries base their data on national accounts, others on satellite accounts or national health accounts. The OECD (Organisation for Economic Co-operation and Development) secretariat aims at creating an international comparable data set using all the information supplied to them based on the sometimes completely different health care systems.

7.2 *History of the health database*

In 1977 the OECD presented an important study on comparative data of the health care branches in its member countries (OECD 1977). Like those of other studies these comparisons were limited to public expenditure on health. These restrictions were caused by the impact of the recessions in the early 1970's following the two oil crises. The recessions caused a break in the medium-term growth perspective. "Lower

rates of economic growth have made it much less acceptable for the proportion of publicly financed social expenditure in the national product to continue to increase." (Abel-Smith & Maynard 1978, page 205). The economic decline highlighted the need to stabilise the share of government spending in terms of a percentage of gross domestic product (GDP), and within government spending the share devoted to health (and social) care.

However, the recession, and its effects, also made it clear that government spending (including social security) was not the whole story. The corresponding amounts on non-government spending were needed to complete the picture, which changed the focus from publicly financed health care to total health care financing and expenditure.

In the early 1980's "... an attempt to provide a quantitative explanation of differences in the pattern of resources allocation to medical care amongst industrialised countries remains a challenge." (OECD 1982, page 1). "Attempts to introduce international guidelines to report expenditure ... of the health systems (with common nomenclatures and estimating methodologies) have not yet succeeded." (OECD 1985, page 17). By constructing its health database, the OECD took on these challenges. In this study 'medical care' is read as 'health care' aiming to create internationally comparable expenditure data.

The database constructed by the OECD in the 1980's differed from previous attempts (e.g. by WHO, World Health Organisation and the World Bank) on four counts (see OECD 1985, page 17):

1. data were collected (or estimated) for all OECD member countries (market-oriented economies). The data collection was not limited to a small selection of countries;
2. data were not collected for a single year but for a relatively long period of time (intended to be two decades, starting in 1960);
3. the database was not constructed by using a specially designed questionnaire, but using existing sources and published journals;
4. aggregates at all levels were not only constructed in money values but in physical values as well (necessary for policy analysis, especially in the area of price and quantity measurement).

Practical considerations dictated that the national accounts be used as a starting point for the conceptual framework on expenditure. The first reason was that ESA (European System of Accounts) and SNA (System of National Accounts) offer a large body of basic rules and definitions common to all countries. The second reason was that the health database to be created was to be part of a larger framework of social policy analysis, which has relations with economic policy. National accounts methodology is a natural link in macro-economic accounting (see OECD 1985, page 19).

The OECD took the SNA as a logical starting point in the compilation of the data for the health database. Acceptance of SNA breakdowns by all member countries is not a sufficient condition to create comparability of the results. One reason is that common accounting standards as proposed by SNA are only gradually implemented by member countries if these standards are seen as useful in the health care field ¹⁾. A second reason is that data are compiled following national institutional requirements, which differ between countries and which make it difficult for statistical agencies to reconcile these data with national accounting standards (see OECD 1977, page 92). So problems concerning comparability of data are not unforeseen, even if the data presented are specifically prepared by statistical agencies for a survey initiated by OECD.

In 1982, the proposed database originally consisted of five sections; in the published version of 1985 (*Measuring health care*) it had expanded to eight, whose relations provided a picture of the existing health care systems in the member countries. These sections ranged from 'Expenditure on health' and 'Health care pricing trends' to 'Social protection' and 'Utilisation of medical services and physical inputs' as well as 'Selected (health status and performance) indicators'. In 1991 the paper versions of the health database were supplemented by an electronic version (a co-operation between OECD for the data and CREDES, Centre de Recherche d'Etude et de Documentation en Economie de la Santé, for the software). In 2001 the electronic version contained, broadly speaking, the same topics collected in ten different chapters, i.e. some restructuring of tables and sections had taken place. In practice the OECD health database has become one of the most widely used 'comparative' tools, which has its merits as well as its shortcomings.

7.3 Purposes

The purpose of the construction of an internationally comparable database on health care and health care expenditure follows logically from the difficulties, which are (see OECD 1990a, page 9):

- data on health care are in general not comparable.
National data are collected for administrative purposes, based on specific national structural features, using different definitions and a non-standardised format for reporting;
- the performance of any health care system is not easy to evaluate.
All countries attempt to supply access to the health care system as effectively and efficiently as possible. Because measuring health care output or health outcomes is extremely difficult, attention focused on inputs, which is also difficult to measure given the lack of standards in terminology;
- it is difficult to measure and control for all kinds of differences between countries.
Differences between countries resulting from economic, social, medical, cultural and demographic circumstances result in "real" differences with regard to

comparability. Of course these differences need to be explained but they should not be treated as part of the problem of data comparability;

- policies are not easily transferred from one country to another.

Although countries can learn from each other's systems and solutions for apparently the same problems, this does not mean that policies that work in one country can be transferred to another. Success in another country would require the same set of incentive structures and behavioural responses from all parties concerned.

Although the OECD publications do not explicitly mention the purposes of the study, it is clear that the basic aim is to "... establish a consistent statistical database which will permit international comparisons ..." on the growth of and access to health services, resource allocation, effectiveness and efficiency (OECD 1985, page 9). In short there is "... a need to know how health systems behave ... as economic, social and administrative entities." (OECD 1985, page 9).

The components mentioned above can be interpreted as the construction of internationally comparable data on the health care sector, as complete as possible with all relevant interrelations, to be used for national goals as well as international comparative studies.

7.4 Approaches to comparability

In theory international comparisons can be broken down into three main approaches:

- an international or other agency fully finances and controls a survey (e.g. those carried out by Abel-Smith for WHO in the 1960's; see References);
- (statistical) agencies from various countries work together to harmonise concepts, classifications and definitions;
- a researcher or analyst uses the available data and "massages" them to make them comparable.

OECD omits to mention the possible combination of elements of the three above approaches.

OECD (OECD 1990b, page 119, 120) notes that the first approach is the surest way to attain comparability, but is costly. However, in my opinion, the results Abel-Smith presented in the 1960's are, although very valuable, not completely comparable. Also in the more recent past results can be shown that are not comparable although they were based on truly harmonised questionnaires (e.g. EU health interview survey data on perceived health).

The second approach is the most classic and is used in the SNA and ESA in the construction of national macro-economic aggregates. This approach may be inevitable in the long run and implementation may indeed require many years. But

it is doubtful whether it will solve the real problem in comparability at a lower level of aggregation, i.e. the disparities in sets of activities performed by the various units of production in various countries. And these sets of activities are not constant in time either.

The third approach is the least expensive and the least reliable, because judgements on the changes needed to create more comparability are made by the compiler only. The OECD health database in the 1980's belonged to the third type. In recent years attempts are introduced to move in the direction of the second approach.

A combination of elements could be very profitable if it were able to combine local expertise of the participating countries, uniformity in data collection and data production (modelling techniques, estimation, etc.) and harmonisation and agreement on the methodologies, classifications and definitions used. This approach tries to combine the advantages of the other three, but consequently many of the disadvantages are combined as well. Without an international legal base or at least strong international pressure for harmonisation, the long-term goal of agreement on cross-country comparability, will hardly be reached.

7.5 *OECD health database and comparability*

7.5.1 *Introduction*

From the start, OECD intended to create comparable time series (for single countries and between countries). Many statistical agencies were not able to establish these long data series because they lacked detailed basic information or resources following changes in nomenclatures and estimating methodologies that occurred in many countries. Another reason for lacking comparability could be the use of series inconsistent with one another, because of different concepts used by different recording agencies in various countries. As early as 1985 OECD noted that "... long-term trends ... are likely to be more accurate than the comparison of level differences." (OECD 1985, page 10). If levels of health care expenditure (or other variables) are less comparable than trends, this means that the contents (in terms of activities performed) of these variables differ between countries. Although the OECD report stated that trends matter more, many authors and official agencies have taken an opposite stance. The ultimate goal in cross-country comparison should be on the level of variables across countries and not on the change in time or the growth rate of a variable in time.

7.5.2 *Multiple sources*

Another problem already noted in 1982 is the use of multiple sources for every country. Multiplicity of national sources creates a multiplicity of incomparability, not only when comparing different countries but also within a country. Statistics Netherlands conducted a study of differences and correspondences in in-patient health care based on two different sources. The study concluded that around

90 percent of the units involved in health care could be retraced to either source, representing roughly 95 percent of the receipts value (CBS 1982, page 14). Although it can be concluded from this Dutch comparison of two sources on the same topic that there is extensive agreement, there is still a difference of some 5 percent (in expenditure) to 10 percent (in entities traced). So, which source is used for international comparisons is quite relevant for the results of the comparison.

Using multiple sources across countries for a single country's database may create a much more complete data set. However, it may also create more problems in cross-country comparability, which cannot always be visualised or traced back to the sources. Lastly similarity of titles is no guarantee that the underlying concepts are identical. Similar entities can play different roles in the health care systems of other (neighbouring) countries.

7.5.3 *Data comparability*

International comparisons are difficult because the data compared differ as there are "... no universally agreed definitions, ... no accepted 'social accounts' definitions of health aggregates, ... [and because] boundaries of health systems differ." (OECD 1987, page 15). In a national context boundaries are seldom important in the analysis of the performance of a country's system. The same is true for "... widely divergent definitions of health facilities or health professionals." (OECD 1982, page 2, OECD 1985, page 17) among the OECD member countries.

Cross-country comparisons are also difficult because of methodological problems and/or conceptual difficulties resulting from the comparison of countries in different stages of economic development, and with different demographic, cultural and institutional structures. Institutional settings determine how health care systems operate in society and are real differences to be dismantled in international comparisons. Cultural attitudes towards health care, economic developments and demography influence the level of expenditure, but are real factors in cross-country comparisons. These real differences should be explained following the comparisons, not equalised or deleted. Lastly, existing accepted definitions and classifications assist comparability across countries but are not sufficient to create comparability. Definitions and classifications do not delete the differences in how activities are performed by the various parts of the health care provider system.

Although comparisons are difficult to make, there is a growing demand for comparable data and their dissemination. In the absence of internationally agreed definitions and concepts, a regular dialogue is present between the member states supplying data and information and the OECD secretariat.

The international comparison of the health care data undertaken by the OECD "... cannot remove all outstanding differences in the countries' health statistics, deeply rooted in their institutions and related reporting mechanisms, but a modest amount of reclassification according to homogeneous concepts has been undertaken." (OECD 1985, page 17). Reclassification means basically that data are

changed from how they are presented by the national sources, leading to different data sets for the same country, the same years and on the same topic. It does not mean, however, that one concept (the international comparable one) is superior to another national concept. It just means that the concepts are different because they are used for different purposes: international comparability and national analysis.

7.5.4 *Metadata and converging concepts*

The term metadata²⁾ is not used as such in the health database. What it does provide are the sources and methods used. In the database concept this means that sources used by the OECD secretariat are listed for every table and if necessary and possible the methods used by the data suppliers to create the data presented. The 'massaging techniques' performed by the secretariat (in the 1980's) are not always presented as such. In a number of cases this is not necessary because they consist of removing or transferring blocks of information supplied by the data providers.

What is perfectly clear, however, is that right from the start of the health database the background information and detailed information on how data are created and supplied for international comparisons is very important. This is true not only for the users of the data but also for the producers, both of the international database and the national data sources.

In the knowledge that the results based on the existing databases are not completely comparable, a conclusion of the OECD health database could be phrased as "Gradually countries are adopting for their domestic health accounts concepts which are internationally similar. Systematic harmonisation is [in 1985] ... probably still a decade or more away." (OECD 1985, page 17). Convergence of health care concepts is progressing slowly, at least in the EU. Systematic harmonisation, however, had still not been reached in 2002.

Closely linked to the above-mentioned conclusion on converging health care concepts is the notion that "... this study [Measuring health care] will probably facilitate the future establishment of an international health accounting framework." (OECD 1985, page 17). With the construction of the System of Health Accounts (SHA), a project executed in close co-operation between the OECD and Eurostat, a first version of this framework came into existence. In this framework instruments that facilitate systematic harmonisation are adopted. Chapter 12 is devoted to the SHA and its possibilities for improving comparability of health care expenditure data.

7.5.5 *Contents of items and comparability*

One of the most important remarks made in several OECD reports (e.g. Measuring health care) concerns the contents of the items in the comparisons. "Between countries similarities of titles does not ensure identity in the underlying concepts." (OECD 1985, page 18). "The most common pitfall consists in using parameters with identical names as if they were comparable. Although similar terms are used the meaning may differ."

(OECD 1990b, page 120). Closely related to this remark is the fact that "... the size of the disparities suggests that more detailed micro-comparisons are needed to understand some of the macro-evidence." (OECD 1985, page 95).

The remark 'like is like' for different country's health care systems will lead policymakers and others to conclusions that are inherently false, because compared between countries, the contents of institutions in terms of activities differ, even though their names as part of the subcategories are identical.

Detailed micro-comparisons can only be carried out if there is detailed information available, not only with respect to the data of the component parts, but also to the contents in terms of activities and accompanying metadata of these parts. Without this information it is impossible to conclude whether parts of subcategories or units in a study are different or comparable.

Contrary to the remark on 'like is like' OECD notes that data are often not comparable because of "... definitional differences, the frequent inclusion in hospital spending of in-patient physician services and the difficulty of separating outpatient hospital spending from total [hospital] expenditure." (OECD 1987, page 60). Inherent in this statement is the basic problem that institutions with the same names are not equal (in terms of activities) in all countries. The basis for the construction of better cross-country comparability is unification of the set of activities to be compared.

Basic characteristics of the composition of the health care providers determine the comparability of health care expenditure, not the surrounding circumstances, like the economy, environment, population etc. These surrounding circumstances determine the health status and influence the level of expenditure. They are real differences where comparability is concerned. Various articles especially mention the topics of the economy, environment and population as determining factors in the lack of comparability (see e.g. OECD 1987, page 15, OECD 1990a, pages 9 and 10). The resulting differences in the data across countries need to be explained, not erased.

7.6 *Final remarks*

The need for reliable and comparable information on health care systems is essential. Experience from international comparisons could assist countries to respond to the problems they face. But "... few ... social programs are framed in an international context ..because of the intrinsically national nature of social systems and the dearth of valid internationally comparable data." (OECD 1990b, page 119). So it may be surprising that an international database on health care exists at all. OECD 1977 concluded that "... in view of the uncertain nature of the data, its non-comparability across countries and the conceptual problems involved, ... the chief usefulness is highlighting issues rather than offering firm conclusions" (OECD

1977, page 86). In my opinion, in 2002, although comparability of health care data across countries of the European Union and the OECD has improved, this basic conclusion is still true.

The conclusions drawn by OECD (and other researchers) from the amount of and the trends in the data rest on the quality of these data. These data are at present still not fully comparable, and this restriction should be borne in mind. Any inadequacy or lack of total comparability in the data is dwarfed by the advances in the area of comparability that can be achieved by using the data.

Looking back on this chapter the questions raised at the beginning cannot be answered in a single unambiguous way.

Comparability of the results is an important issue in the OECD health database although it is not present to the full extent needed. The basis – SNA methodologies and classifications – are not used by all member countries. On the one hand SNA does not guarantee comparability, on the other alternative systems might perform much better from a comparability point of view. The data supplied by the member states is rearranged (or ‘massaged’) by the OECD. This means that results can be improved, in spite of the use of different reporting systems.

Consensus on concepts, boundaries, definitions and calculation and estimation methods has been reached in few areas of health statistics. In the OECD health database, tables require “... a page of small-print footnotes to explain the departures from the ‘desirable’ concepts.” (OECD 1993, page 41). The problems are enlarged by the heterogeneity of sources used in the member countries and the tendency of many national statistical agencies and government bodies to use domestic boundaries and nomenclatures, even when international guidelines exist. This last point is not surprising, because often SNA is not applied in health care statistics because of its lack of flexibility for example in the use of classifications and definitions. Nationally oriented systems of health care statistics are much more usable. So, although the starting points of the OECD health database are clearly found in the SNA framework, this is obviously not followed by all countries supplying information to the OECD. This ‘opposition’ of countries to the use of SNA as a standard for compiling health care data has, however, apparently not been detrimental to the comparability of the data collected and supplied.

Since the beginning of the OECD health database, the level of comparability has increased, but the process remains long and gradual. The experience gathered from other comparative data systems (like SNA) suggests that several ‘vintages’ are needed to attain a reasonable level of satisfactory comparability. After several vintages, it was noted by the OECD in 2001 that although there is “... growing agreement about international definitions of health variables, and [a] growing adherence to these definitions ... there remain many definitional divergences ...” (OECD 2001, page 9). These differences are shown in underestimation of the data supplied, health care boundary problems between countries, quality of private

health care expenditure (estimates) and frequently reported breaks in series (see e.g. OECD 2001, page 40).

However, after two decades of health database information gathering by the OECD, it seems that the possibilities of reaching a substantially higher level of comparability, using existing tools, are diminishing. Cross-country comparability is hampered by the fact that actors in the field of health care perform different sets of activities in different countries, and this cannot be redressed by improved definitions. Cross-country comparability can only be improved further by comparing identical activities and sets of activities.

Nevertheless it can be concluded that the OECD health database is not only one of the most widely used tools in comparative studies, but that it also created the basis for the attention for comparisons of health care data for international use. This fact contributes to a large extent to the development of the System of Health Accounts (SHA, see chapter 12).

Notes:

- 1) Just look at the terms production value central in SNA/ESA, and expenditure central in health accounting. In SNA terms the difference between these two terms is that production value is a term used in supply and hence measured in basic prices, while expenditure is a term used in uses (of the products) and hence measured in purchaser's prices. See also note 7 in Chapter 4).
- 2) See chapter 2 Section 2.2.1 for a discussion on metadata in health care.

8. *The statistical process*

8.0 *Questions*

Do the various ways data are collected and treated within the statistical process affect cross-country comparability of health care, and if so can this effect be changed so that it is positive for comparability of health care data?

8.1 *Introduction*

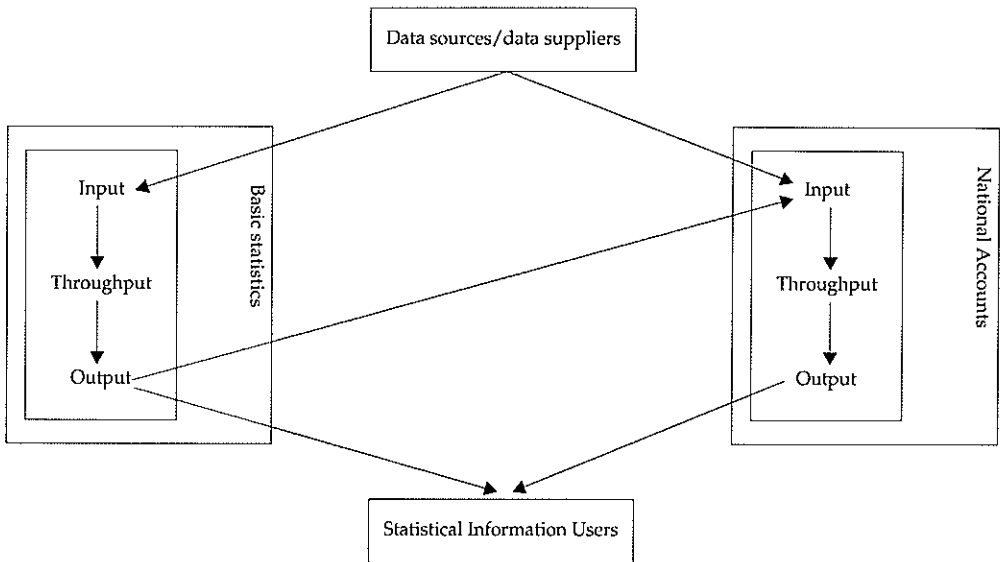
The previous chapters have discussed existing frameworks and classifications and their relevance to comparability. These chapters did not go into the statistical process as a logical step in creating the statistical data, the systems and the methods used. This chapter examines the statistical process, with special attention (if necessary) for health care.

Before any data can be presented or compared, a complicated process has taken place which can be broken down into three main parts: the input acquisition process, the aggregation process, and the output delivery process (see e.g. Sundgren 1993, page 5). In my opinion, Sundgren's 'aggregation process' is an insufficient term for the middle part of the statistical process; this part of the process encompasses more than just aggregating data. Statistical production or throughput process would cover it better.

The purpose for which information is gathered influences the process. Statistical information is gathered and used for statistical purposes. Other purposes (e.g. administrative) may require other types of information or a different treatment of the same type of information.

The statistical process of statistical authorities, who supply information for public use, can be characterised by six leading principles¹⁾: impartiality, reliability, relevance, efficiency, administrative safeguarding and transparency. The information supplied needs to be impartial and reliable, i.e. the results are not influenced by interested parties and are created in a professional way. The data need to be relevant for society and produced efficiently (in terms of costs and response burden). Confidentiality of individual data is guaranteed, so that data suppliers are prepared and stay prepared to provide data that are true. Lastly transparency relates to the decision processes regarding statistics to be produced and the way in which the production process is executed. All these principles are central in the production of statistical information that is indispensable for society and science.

8.1 Statistical process



The statistical process – input, throughput, output – is not only relevant for basic statistical processes describing a single activity or a limited number of phenomena, but also for (partially) integrated systems like health care accounts and national accounts. The difference being not only that integrated systems like national accounts integrate a lot of statistical information into one system, but also that the inputs for national accounts are (largely) end-products of basic statistical processes. For a complete picture of this process see figure 8.1.

On the input side data sources and data suppliers are essential, while on the output side the users of the statistical information are the final goal for which all the processes were executed. Between input and output, we have the throughput. The demand for statistical information and the demand-linked user wishes need to be taken into account in constructing and maintaining a statistical programme. Apart from technical situations influencing the process, the human factor will also influence the way the statistical process is executed. The three parts of the statistical process will be described and discussed below with special emphasis on cross-country comparability of data.

8.2 *Input process*

The whole input process rests on the core element of availability of data. Essential characteristics in the input process are the creation and maintenance of a database of

survey units and the creation and sending out of questionnaires (either on paper or electronically). If sampling from the basic register is required, sampling techniques are also important. The next steps in this process are the actual collection of the data and the data entry process. Other characteristics of data availability have to do with the response burden for companies, corporate accounting systems and cultural differences within and between countries.

– *Registers and data items*

It goes without saying that the units included in or excluded from the register to be used as a survey frame can and will influence the results. Just as important are the items selected to be included in the questionnaires, as well as the way these items are described (metadata on the data collection). Different starting points in how these topics are handled in a country may restrict the possibilities of adding up national results for different topics. Adding different countries in this process influences the cross-country comparability of the topics concerned.

Although sampling techniques are determinant for the input process, their results are relevant in the throughput process.

– *Data collection methods*

Data can be collected by direct methods (from the respondents) or indirect methods (from respondents through intermediary). Data sources used to be limited to questionnaires – sent out by statistical offices and completed and returned by respondents – supplemented with data from published accounts and reports. Today it is increasingly usual to tap existing administrative sources, both directly and indirectly. Especially when indirect sources of information are used it is imperative to know how the data have been processed by the holder, just as the statistical office would check its own procedures in the construction of statistical inputs.

– *Accounting systems*

A very important factor on the input side is the use of different accounting systems. Even within one country the various accounting systems cause problems in internal consistency and comparability, problems that are enlarged when data are compared between countries. Differences in concepts constitute another important obstacle. Concepts used in administrative systems and developed for administrative uses almost always differ from concepts used in national statistics. Corporate administrative systems are not easily changed, they are ‘corporate bound’. Changes can be introduced if they are important for the company or at least do not contradict company regulations.

On the other hand the existing problems for comparability will not be solved by such savings in the daily operation of the statistical production process, as is clearly stated in the French contribution to the Conference of European Statisticians held in Paris in June 2000 in the framework of the United Nations Statistical Commission and Economic Commission for Europe. “The conflict of interest between the benefits of using administrative sources and the need for

greater harmonisation of statistical systems and statistical output needs to be resolved. Harmonising concepts is a necessary precondition. Harmonised surveys run counter to the objective of reducing the burden on the respondent and do not provide ipso facto more homogeneous results. Harmonised questions do not guarantee harmonised answers. Responses to questionnaires still refer to national concepts and conventions. Genuine statistical harmonisation should therefore presuppose as a prerequisite homogenisation of non-statistical concepts. It is necessary to define concepts and classifications that have a satisfactory level of universality." (UN 2000c, page 8).

– *Financial resources and constraints*

Another important factor, not mentioned by Eurostat but greatly influencing the possibilities of official statistical offices, is the existence of the European Monetary Union (EMU), more precisely the consequences for national government policies following the difficult adaptation process in trying to comply with the restraining conditions of the EMU stability pact. For almost all European Union (EU) countries these requirements mean cutting back government spending to keep government deficits within agreed limits. Even in cases where statistical offices are not part of the public arena, the consequences are certainly felt, because less government spending will lead to less public information. For statistics the surrounding external conditions have certainly changed, affecting the quantity and/or quality of data and thus cross-country comparability.

On the other hand, the EMU also means that European organisations like the European Central Bank need more (statistical) information (e.g. on short-term indicators), a need that can easily conflict with the more limited resources available in the countries of the European Union.

– *Response burden*

Related to this topic is the wish to reduce the burden placed by statistical surveys on enterprises. In some countries (including the Netherlands) this wish has been expressed and measures are being taken, resulting in a reduction of statistical information collection. For statistical offices the demand for a smaller burden on the one hand, and more international information on the other, means at best a status quo of the existing number and size of surveys. Limiting the amount of information gathered on specific topics is detrimental for cross-country comparability.

– *Attitude towards information*

An important factor, but one very difficult to determine, is the attitude of society towards the supply, demand and handling of information. In the USA great importance is attached to record filling and the quality of the data. In many European countries this task is delegated to the lowest possible level.

The more open and co-operative a society's attitude is, the more likely it will be that enough data are collected that are in principle of high quality. High quality national data are advantageous to cross-country comparability. It is well-known, however, that in some countries the eagerness to comply with requests for

information is diminishing among some parts of the population and in some segments of industry.

This cultural factor will not only influence the quality of the national data but also the quality of internationally requested data and cross-country comparability. In combination with the generally formulated wish to reduce the response burden, at least in the Netherlands, statistical authorities increasingly use registration data and data collected by third parties (like umbrella organisations) in the construction of statistical information.

8.3 *Throughput process*

The core activity in the throughput process is the processing of the data. Characteristic factors in this process are estimating techniques linked to the sampling process in the input phase, error detection and correction techniques and lastly characteristics of the data themselves, like consistency, completeness, coherence and continuity. Just as important as the actual data handling and analysis are the classifications used in the processing. And although hardly ever mentioned, the human factor is decisive in the statistical process. All these factors will be discussed below in so far as they influence the quality of the results and thus cross-country comparability.

– *Sampling and estimation techniques*

Sampling techniques can greatly influence data quality, just as the processes used in the actual collection and storage of data. Although methodologically correct, samples and corresponding estimation techniques (correcting for non-response) may ultimately result in statistically correct data which are still not – in detail – usable for cross-country comparability;

– *Consistency, coherence, completeness, continuity*

In the throughput process components like the integration of the information in a consistent system is important, as well as checks on completeness and additional calculations. Consistent and integrated systems require rules on harmonisation of definitions and concepts as well as the use of agreed classifications. However, this is not enough. Without knowledge and skills (the human factor), consistency and continuity in statistics is hard to maintain;

– *Classifications and standardisation*

On the topic of classifications and standardisation it is noted (as recently as 1996) that “There is no acceptance of the fact that standardisation is needed of data collection and reporting methodologies as cross-country comparisons may lead to misleading results according as which components of health expenditures are included in different countries.” “In order to carry out an international comparison of national health care expenditures a standardised definition of

health care expenditures is needed. However measurement of health expenditures in different countries varies considerably as different methodologies are usually applied. Similar problems pose the different definitions as well as the different accounting systems within a largely homogeneous entity like the European Union." (Kanavos 1996, page 15).

A standardised definition of health care and health care expenditure could be very helpful in tackling the problem of cross-country comparability of health care. However, not only are the definitions and components used in various countries determining factors in comparability. The use of exactly the same components does not guarantee comparability across countries because the content of these components varies and people influence the use of methodologies and classifications and the interpretation of these tools;

– *Analysis and the human factor*

The human factor is decisive for the statistical process: people handle, interpret and use the tools at their disposal to get the best results possible.

Even if all the technical factors influence the statistical process in the same way in all countries, the construction of statistics and national accounts is still human work, relying on the expertise and knowledge of individuals, as well as on how they interpret classifications and definitions. So, another important factor influencing comparability relates to the skills and interpretation of the analysts, not only in the national accounts departments, but also in the departments creating the basic statistics and other documents (administrative records) to be used by the national accounts departments in the integration process. To achieve better comparability, international consensus in judgements of data and systems used reached through international co-operation and consultation would be very useful.

An example is the level of detail needed to fill the homogeneous units of production with the data required to be useful for statistical accounts. For purposes of comparison, homogeneous units of production are superior to kind-of-activity units. For many companies in many countries it is an arduous enough task to complete the questionnaires or supply the data sets needed to create complete kind-of-activity units, let alone the units of homogeneous production. In practice this means that either the company administrations or the statistical offices' basic statistical departments have to make the decisions. For company administrations, the easiest decision is not to fill in what they don't have, a decision that could give rise to discussion with the statistical office that needs the information. If companies are not able to supply the information to the statistical office, the essential blanks have to be filled or have to be replaced by information that can be supplied by companies. On the other hand, company administrations might decide to 'calculate the missing data' in a more or less consistent way, hoping the statistical office will be satisfied with the result. Although in that case statistical offices get all the information they request, the value in the reality of everyday economic life may be unreliable. Whichever

procedure is followed by the data supplying units or the data processing units, comparability based on factual data loses out. Not just because decisions are taken, but because the process and the reasons for the decisions are not documented (at the national business group or class level, or ideally at the level of the individual registering unit) and therefore cannot be traced as sources of incomparability.

8.4 *Output process*

The core characteristic of the output process is the presentation of the results. The way the information is dispersed is one of the most important components. Other factors influencing the output process are the points of view of the users, what information do they need on which topics using which classification instruments, and what level of aggregation is used in the presentation of the data. A last item in the output process is the statistical authorities' level of independence from government or political influence.

– *Choices and user wishes*

An important point made by Eurostat regarding comparability of the results concerns the fact that all "... attributes of official statistics: accuracy, rapidity, timeliness, comparability in time, comparability among countries, detail, consistency, coherence, reliability, adequacy, relevance, transparency, handiness, or openness [...] can be traded not only against costs, but also against each other, a little bit more of one against a little bit less of another." (UN 2000b, page 1). Different users may require different choices in attributes. Unless the statistical authorities have limitless resources, it is hardly possible to fulfil all wishes of all users. Choices in attributes are determined largely by the importance of the users of the statistics. Cross-country comparability of parts of national accounts will be high on the political agenda in the European Union, because it has financial consequences. Health care and cross-country comparability of health care data are less important within the framework of the European Union and will be lower on the list of priorities and work to be done.

– *Level of aggregation*

A last factor influencing comparability across countries to a large extent is the content of the area to be compared. As soon as the macro-economic level of national accounts is exchanged for a another level of aggregation, the boundaries of the area compared play a decisive role. The content of the thematic areas compared across countries are not the same in the various countries.

As the level of detail increases – from the more global macro-level of SNA to the less aggregated, single policy area like health care (or even the very detailed distinctions that can be made within this area) – the tools available for comparative analysis become more numerous, and the level of insight into and

accessibility of the description of the phenomena increase. However, the problem of comparability is not solved. As the level of detail increases the differences in the contents of the phenomena – that are described across the various countries – are clarified. Comparability profits from an increase in harmonisation of the contents of the phenomena in terms of activities.

– *Level of independence*

Just as the attitude of the population towards information influences the input quality of the data, the level of independence of the statistical authorities responsible for official statistics greatly influences the statistical process and thus the comparability of the data. Statistics Netherlands, for example, is completely independent and has no economic or political interest in the results it produces. A factor related to the level of independence is administrative safeguarding, i.e. suppliers of information must feel certain that the fact of their supplying this information will have no consequences, either positive or negative, for them on an individual basis.

In other countries statistical authorities are sometimes less independent from political organisations like ministries, in respect of handling and presentation of data. Cross-country comparability may be harmed by these different attitudes towards independent statistics in various countries.

8.5 *Final remarks*

The basic steps in the statistical process are input, throughput and output. The input is largely dependent on data available in the outside world and often beyond control of the statistical authorities. The output, although determined by the statisticians, is dependent on the demands of users often resulting in differing sets of information.

The statistical production process itself is under control of the statistical authority, but boundaries are set by the legal system in which the office operates. Within these boundaries statistical attributes and processes are somewhat interchangeable. Cross-country comparability (especially of such a detailed area like health care) is not high on the priority list.

As a result of financial and other constraints, and perhaps positively influenced by technical possibilities, more and more use will be made of individual and collectively processed large databases to fulfil the data need of the statisticians. Tapping data from individual firms or collective databases offer possibilities, but without harmonisation of definitions and classifications, comparability will not be improved.

At the macro-economic level only the boundary of the economically relevant process is important when comparing across countries. In comparing the aggregates of national accounts across countries it is noted that “The methodology applied by different countries in the measurement of GDP may differ considerably, and this

may be due to the methodologies according to which data are collected, compiled and analysed, or historical reasons. "Consequently the basis of reporting data across countries may be entirely different." (Kanavos 1996, page 15).

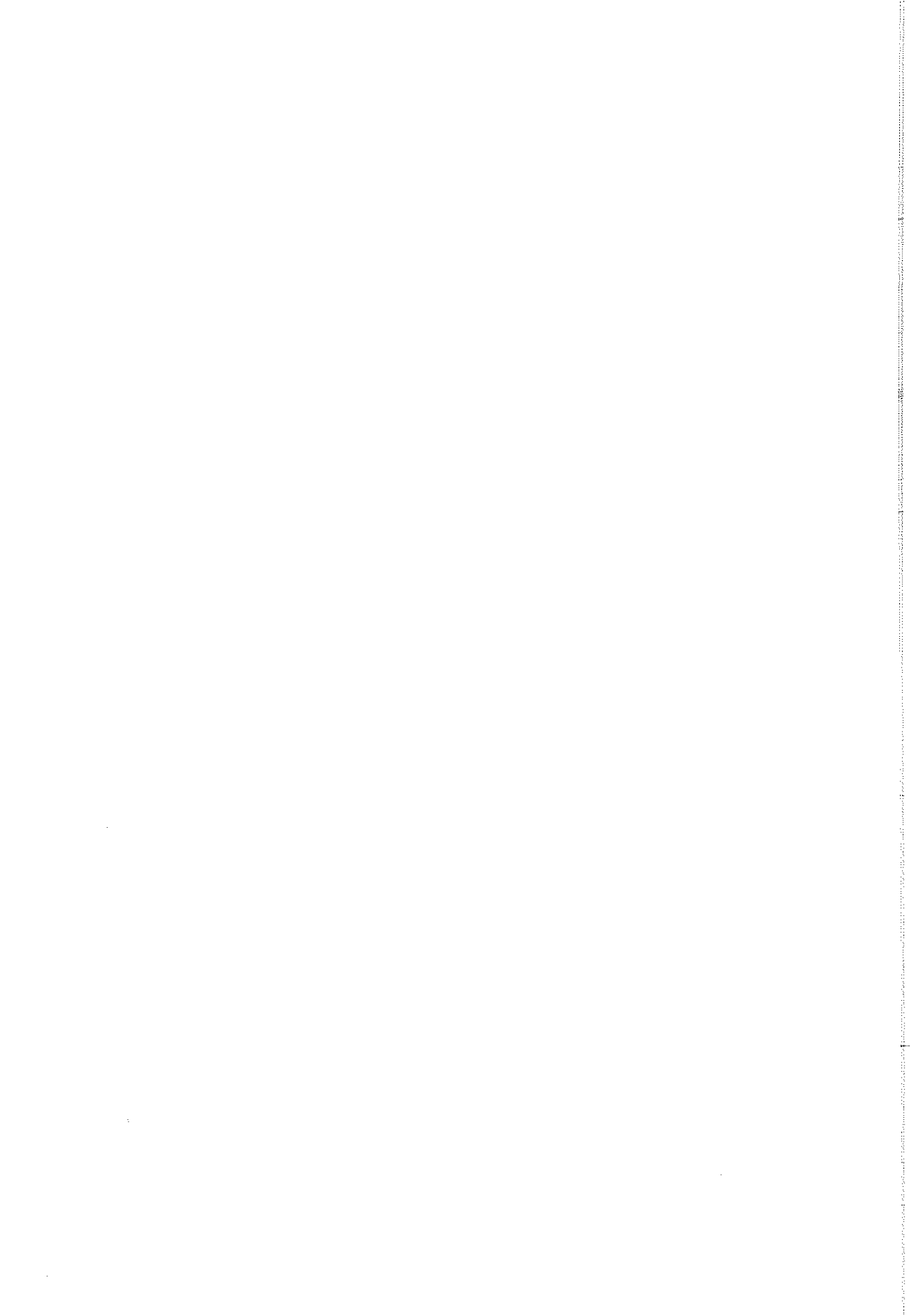
Leaving the national accounts aggregates and going into detail, not only the various distinguishable boundaries become important but also the contents. The contents of the detailed areas to be compared across countries are not identical and this hampers comparability. Cross-country comparability in health care would benefit from an increase in harmonisation of the actors in terms of activities. Uniformity of actors and activities would be even better and is the ultimate goal (that will probably never be achieved).

Each step in the statistical process – input, throughput and output – as well as each step in relations between statistical offices and data suppliers and users can be and in fact are determining factors in the level of comparability of specific statistical information. Statistical authorities can control the internal process of producing statistics: technical processes can be completely controlled, the human factor in the process can only be guided in the right direction. But statistical authorities have less or no control over the outside world, on both the input and the output sides.

All these factors can have positive as well as negative effects on the level of comparability of data. The final result of this conglomerate of factors is, however, undetermined.

Note:

- 1) These six principles are published in article 10 of Council Regulation (EU) No. 322/97 of February 17, 1997.



9. *Comparability of existing data: a practical example*

9.0 *Questions*

Are data on health and social care – in terms of a percentage of GDP – presented by OECD and Eurostat (based on SNA/ESA) for a selection of countries comparable?

Are other measures that can be calculated on an aggregate level (based on SNA/ESA) in concordance with the share in GDP and across countries?

Are the data for the various states of the USA better comparable because they are based on more unified sources and methods, for example?

Are the data in the OECD health database – in terms of a percentage of GDP – comparable?

Has the degree of comparability of a set of data elements increased or decreased within one country over time (time comparability; e.g. in the Netherlands in the period 1953–1999)?

9.1 *Introduction*

Chapter 3 (on the System of National Accounts, SNA and the European System of Accounts, ESA) and chapter 4 on satellite accounts discussed the systems themselves in the light of comparability, and in chapters 5 to 7 the possibilities of functional classifications, health accounts and the OECD health database were presented. In this chapter data as published by the various organisations (Eurostat, OECD and the Centers for Medicare & Medicaid Services in the USA) are presented in order to illustrate the problems involved in comparing data on expenditure and production value in the health area.

There are various ways to present the data. Here we have chosen to present data collected by various organisations, based on distinct frameworks, for a single variable: production value¹⁾ as a percentage of GDP (Gross Domestic Product) for data sources based on national accounts. For the OECD health database, one of the sources used in this chapter, production value is not available, so expenditure²⁾ is used. The choice for the variable is explained below.

This chapter does not intend to explain the differences between production value and expenditure in comparability of health care. The correspondences and differences between production value and expenditure are highlighted in chapter 16 using the Netherlands as an example.

9.2 Data sources on health and social care

To illustrate the problems in comparability of health care expenditure and production value, in view of the known pros and cons of various absolute measures, the data are best presented as shares of a comparable entity. In theory gross domestic product (GDP), gross domestic income (GDI) and gross domestic expenditure (GDE) are the same. Each of these three aggregates could be used as a denominator. In theory it is possible to use GDP as a denominator for production value and GDE

Table 9.1
Health and social care as a percentage of GDP in EU Member States by source, 1997

	SNA/ESA based data		OECD Health Data Base
	EU New Cronos	OECD Service Statistics	
	<i>% of GDP</i>		
EU Average	6,4	5,6	8,7
Luxembourg	4,3	4,3	5,9
United Kingdom	6,6	5,8	6,7
Ireland			6,9
Spain			7,0
Finland	8,1	7,2	7,3
Portugal	5,4		7,5
Austria	4,5	4,5	7,9
Sweden			8,1
Denmark	10,3	8,7	8,2
Italy	4,6	4,3	8,3
Belgium	6,1	5,7	8,6
the Netherlands	7,2	6,5	8,7
Greece		4,6	8,7
France		6,1	9,4
Germany	6,1	5,7	10,5

Sources:

Eurostat New Cronos, tables on National Accounts, 2001: GDP, production value health and social care and ECU exchange rates.

OECD Service Statistics on Value Added and Employment, 2000 Edition, OECD, Paris, 2000: GDP, value added on health and social care.

OECD Health Data Base 2001 Edition, OECD Paris, 2001: Health data and exchange rates.

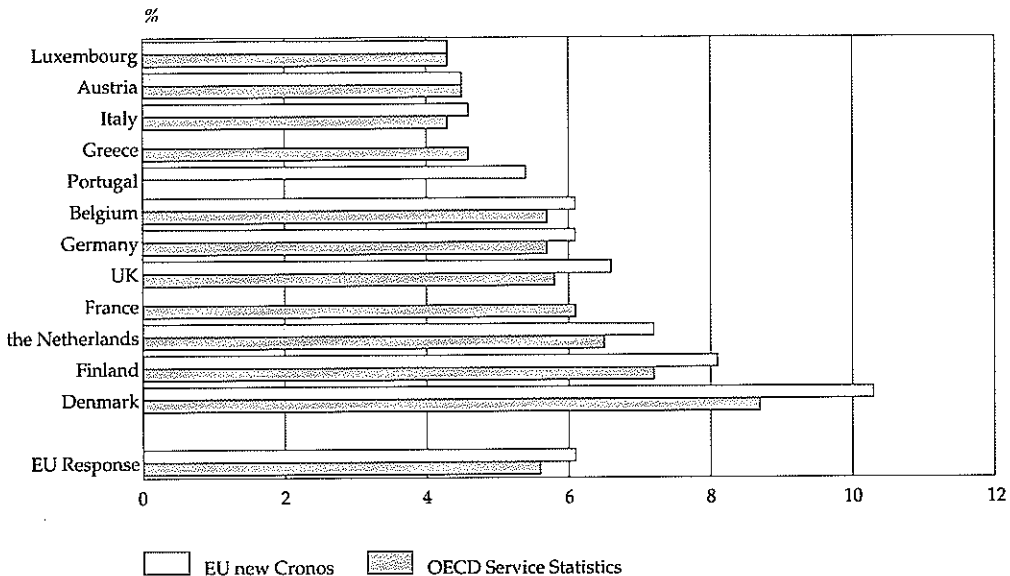
for expenditure data. For pragmatic reasons (data availability), and because it is conventional, GDP is used as a denominator.

To measure the level of comparability, the value of health care and health care expenditure is presented as a percentage of GDP.

The data collected and presented by international organisations are shown for the European Union (EU) member states. Table 9.1 and graph 9.1 compare the data based on SNA/ESA principles (OECD services: statistics on value added and Eurostat New Cronos) with each other and with the data presented in the OECD health database. The EU average presented is a weighted average in which GDP data (measured in ECU for New Cronos data and in US \$ for OECD data) are used as weights.

As health and social care constitutes part of the national economy, it is included in the production value of a country. Indeed, to a certain extent, it determines the size and development of GDP. Some researchers state that taking into account this relationship would require the calculation of health and social care as a share of GDP excluding health and social care (see e.g. Kanavos 1996, page 14). Although it goes without saying that health and social care is an important branch in the totality of GDP, in my opinion this point of view is valid if the branch is considered as creating expenditure. However, from a production point of view every branch or industry

9.1 Health and social care as a percentage of GDP in EU Member States by source, 1997



Sources:

Eurostat New Cronos, tables on National Accounts, 2001: GDP, production value health and social care and ECU exchange rates.

OECD Service Statistics on Value Added and Employment, 2000 Edition, OECD, Paris, 2000: GDP, value added on health and social care.

OECD Health Data Base 2001 Edition, OECD Paris, 2001: Health data and exchange rates.

generates value added, and health care is no different from other industries. In graph 9.2 the shares of health and social care are expressed as a share of GDP, and as a share of GDP excluding health and social care expenditure.

Comparing the data based on the principles of ESA and SNA – which include data supplied by Eurostat New Cronos and OECD services: statistics on value added – the following observations can be made.

Because both publications are based on national accounts principles, the data presented cover the area of health and social care (as included in ISIC 85 or NACE class N). Only 10 of the 15 EU member states report to the EU, 11 to the OECD.

The average share spent on health and social care according to the EU New Cronos (6.4 percent of GDP) is higher than the average reported by the OECD services statistics (5.6 percent). It might be expected that as they are all based on national accounts, data supplied to the EU and the OECD would be more or less identical. In reality only two countries – Austria and Luxembourg – report almost exactly the same data to both Eurostat and the OECD. All the other countries report higher shares to the EU. It is obvious that the difference reported by Denmark (EU: 10.3 percent; OECD 8.7 percent) was a significant contributor to this difference.

Because the classifications and definitions used in SNA and ESA are to a large extent identical, it is hard to believe that the data supplied by the various statistical authorities of EU countries are comparable. It is remarkable that differences in share of GDP between Italy and Denmark, both economically highly developed countries, is so large. The low position of Luxembourg – also an economically prosperous country – can largely be explained by the position of its financial markets. Because of the relatively large contribution of bank services to GDP the level of production value in health care as a share in GDP is relatively low. This is reflected in all comparisons which include Luxembourg, not just for health care.

With respect to the third source – the OECD health database – and comparing this source with the sources based on SNA principles the following remarks can be made.

Although the health database was created on a voluntary basis, all 15 EU member states supply data on health care expenditure. As in eight other countries, the calculated share of GDP for the Netherlands is higher in the OECD health database than in national accounts data.

Because of boundary differences of course, the area considered differs between the two sources. In the OECD health database the area taken into account includes retail sales of medical goods, administrative expenditure concerning health care and health care goods and services produced by government. The SNA includes the health care branch, but also data on social care, which is not or not completely included in the OECD health database. With this in mind, the lower shares supplied by two countries (Finland and Denmark) in the OECD health database become more difficult to explain.

Another difference between these two major sources is the valuation of the subject. In SNA/ESA production value is measured according to agreed rules. In the OECD health database, health care expenditure equals the amounts that actors in the health care branch receive for their services.

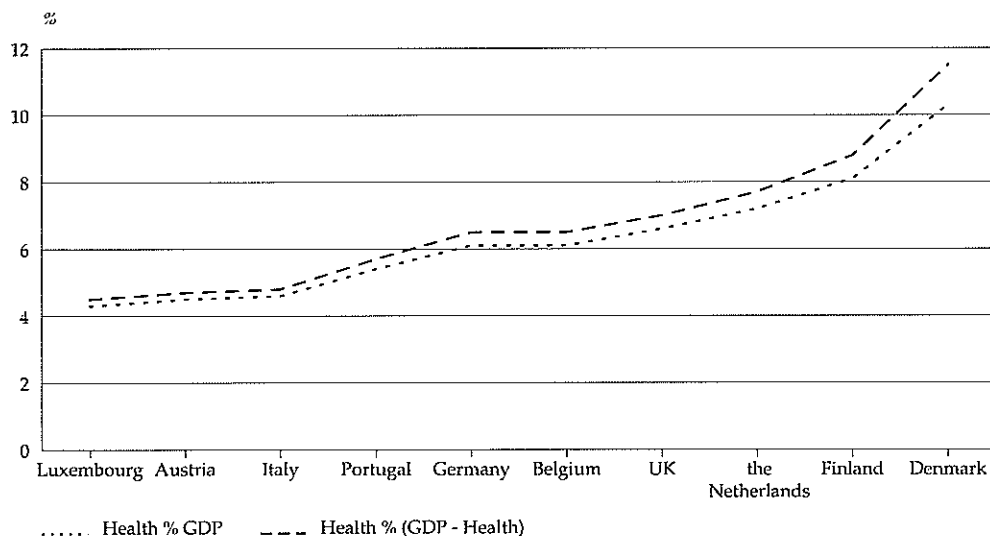
The most important differences, at least for the Netherlands, are connected with the treatment of interest receipts and wage subsidies (and other non-product subsidies), which are not part of production value but are part of the receipts of the producers of health care.

Looking at the complete picture it is difficult to imagine that the data as supplied by Denmark and Germany in these sources are based on the same set of definitions, contents and classifications. The share for Denmark in the OECD health database is 2.1 percentage points smaller than reported for the EU. Germany on the other hand reports a 4.4 percentage points higher share in the OECD health database.

Health and social care as a share of GDP without incorporating the production value of the health and social care branch is illustrated in graph 9.2, based on the data supplied by Eurostat for the EU member states.

Although the shares calculated in GDP excluding health care differ from those in GDP including health care, the general picture is the same. It does not seem to matter whether health and social care is expressed as a share of GDP including or excluding health and social care itself.

9.2 Health and social care as a percentage of GDP (including and excluding health and social care) in EU Member States, 1997



Source: SNA / ESA data: Production value and ECU exchange rates: Eurostat New Cronos, tables on National Accounts, 2001.

9.3 Health and social care shares in SNA/ESA aggregates

Expressing health and social care as a percentage of GDP is only one way of using on national economic aggregates. Calculating other ratios like the share of production of ISIC 85/NACE N in total production of the EU countries, or the share of consumption expenditure on health and social care in total consumption expenditure provides additional insight. Table 9.2 gives some data for a selection of EU member states.

Table 9.2
Share of NACE class N in various National Accounts aggregates, 1997

	% of GDP	% of Production ¹⁾	% of Compensation ²⁾	% of Consumption ³⁾
Luxembourg	4,3	2,9	5,4	n.a.
Austria	4,5	4,0	5,1	3,1
Italy	4,6	4,5	7,3	3,3
Portugal	5,4	4,8	6,9	6,9
Germany	6,1	4,9	8,0	4,0
Belgium	6,1	4,3	7,7	4,0
United Kingdom	6,6	7,2	8,6	n.a.
the Netherlands	7,2	4,8	9,4	4,0
Finland	8,1	5,6	13,4	3,5
Denmark	10,3	7,2	14,7	2,4
EU Average	6,1	5,4	8,3	3,8

¹⁾ Production value is the value of the goods and services produced measured at market prices seen from a supply or supplier's point of view.

²⁾ Compensation of employees is defined as the total remuneration, in cash or in kind, payable by an employer in return for work done by the latter during the accounting period.

³⁾ Final consumption expenditure consists of expenditure incurred by resident institutional units on goods and services that are used for the direct satisfaction of individual or collective needs.

Source: Eurostat New Cronos, tables on National Accounts, 2001.

In three of the four macro-economic aggregates (share in GDP, share in production and share in compensation) presented in table 9.2 the ranking of the countries is more or less the same. In the fourth economic indicator, on private consumption expenditures, the positions of both Portugal and Denmark seem odd. Denmark ranks lowest in final consumption expenditure and Portugal highest, exactly the opposite of the rankings on all other aggregates. Portugal appears to have one of the highest cost-sharing ratios of the EU for health care.

While one might expect the relative differences between the macroeconomic variables across countries to be within rather a small range, the data show some

anomalies. For five countries the share of NACE N in production is around 30 percent lower than the share in GDP. However, in the UK this share is 10 percent higher. Comparing the shares in production and the shares in compensation of employees, it is striking that both the UK and Denmark show a share in production of 7.2 percent. In compensation of employees, however, the UK share is 20 percent higher. In Denmark the share in compensation is more than 100 percent higher than production, which is virtually impossible.

All the data presented above are derived from national statistical authorities, i.e. statistical offices, ministries of health and other national bodies. National sources (even those based on SNA/ESA) have their own national intricacies and ways of handling data for statistical purposes (see chapter 8). Below we look at the USA, where all the independent states in the federation use the same classifications and methodologies in the production of the data and in the supply of information to national offices on specific topics. Health care is one of the topics for which detailed information is available.

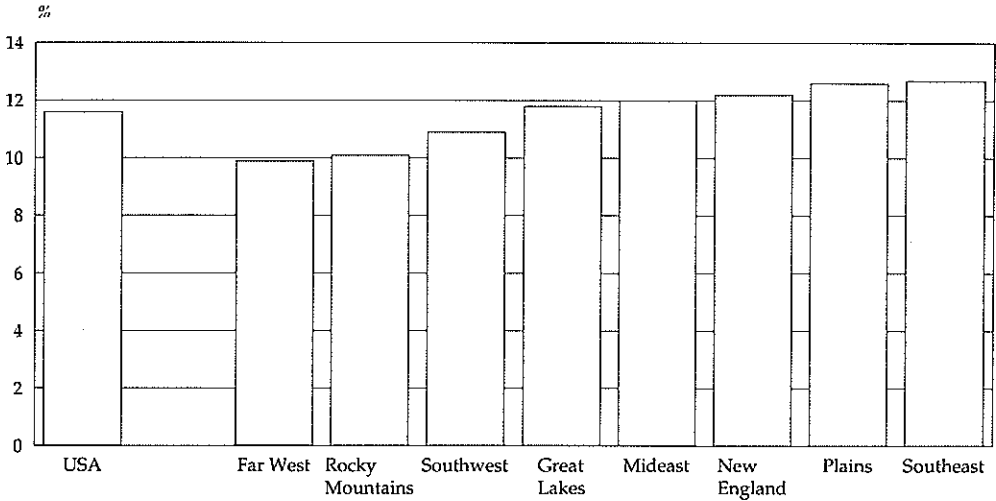
9.4 *Health care data in national health accounts of the USA*

Chapter 6, on health accounts, already mentions that cross-country comparability is made difficult by differences in the health care system. Especially the boundary problems concerning the division between health care and social care on the one hand, and health care and health care related items on the other, distorts the comparability of the data. In the USA the boundaries of health care in relation to social care and in relation to health related items are set uniformly for all states, as is the method of estimating the data. As health care systems across the USA are not unified, these differences still influence the level of the data.

In this chapter, and in the chapters 3 and 5, it is stated that using SNA/ESA methodologies does not guarantee comparability of the results on health care expenditures either. One of the sources of incomparability of the EU and OECD results – interpretation of the methodologies and differences in classification use – can be eliminated by using USA data. In the USA one organisation is responsible for the statistical production of health care data using one single method and one classification. With such common input and throughput components, it might be expected that the spread in the results between the states of the USA is smaller than the spread in the EU and OECD member states.

In graph 9.3 the shares of personal health care (PHC) in total gross regional product (GRP) for 1998 are presented as recorded by CMS. Gross regional product is the sum of the gross state product (GSP) of the states belonging to the region. PHC is defined as the total of health care expenditure minus the expenditure on public health programmes and activities conducted by the government, administrative

9.3 Personal health care percentage of GRP ¹⁾ in the USA, 1998



¹⁾ GRP: Gross Regional Product: Sum of the Gross State Product (GSP) of all the states belonging to the designated region.

Source: Health Care Financing Review, Summer 2001, table 4.

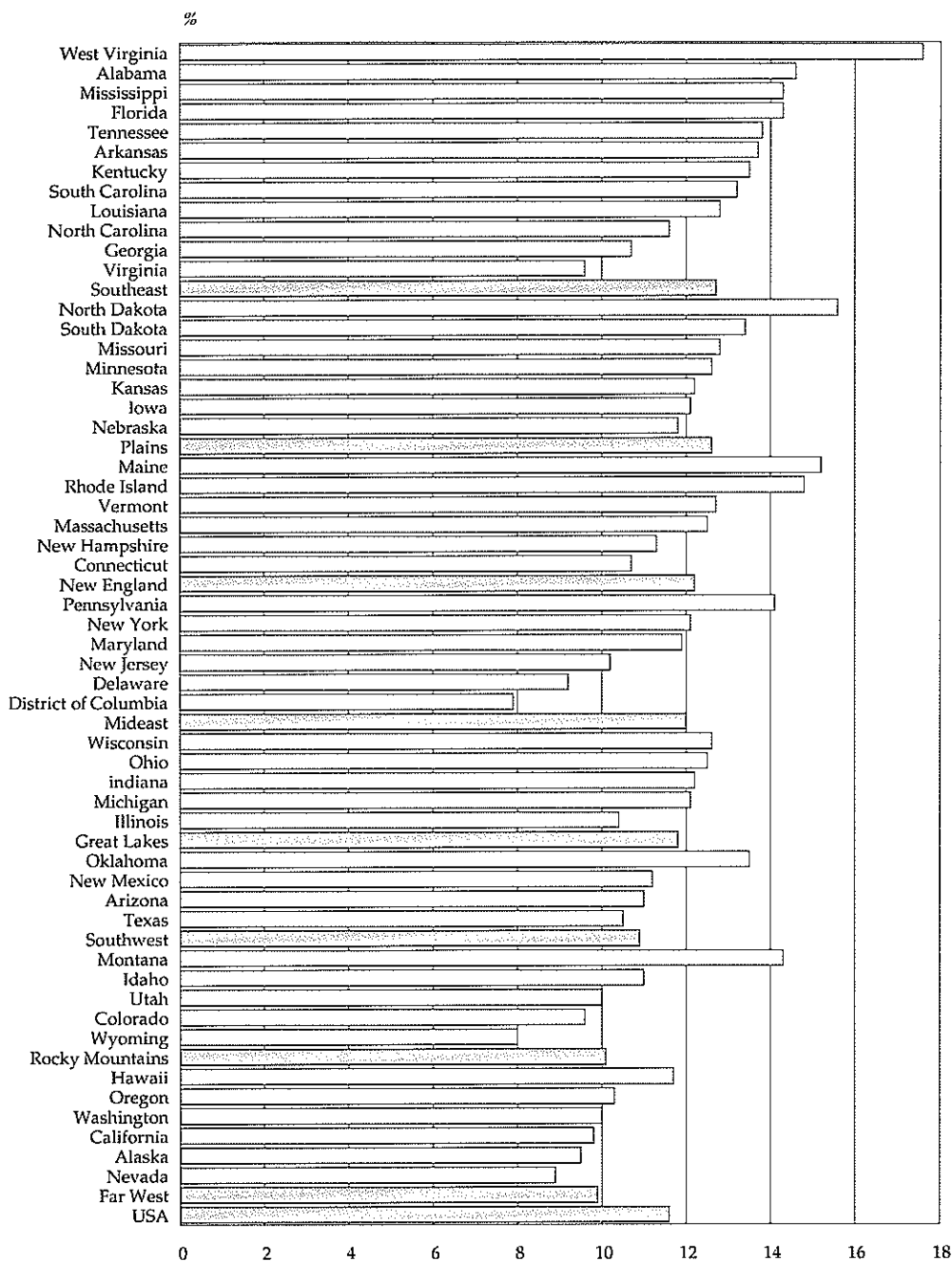
expenditure by government and private health care insurance companies, and lastly expenditure on investment in health care.

The difference in share of GRP between the lowest region (Far West, 9.9 percent) and the highest (Southeast, 12.9 percent) is 3.0 percentage points, far smaller than the difference recorded for the EU countries (around 5 percentage points) in the OECD health database for the same variable (PHC) or in New Cronos on health and social care (around 6 percentage points).

However, looking at the same data for all states in the USA (GSP) the difference between the lowest (District of Columbia, 7.9 percent) and the highest share (West Virginia, 17.6 percent) is larger than the difference recorded for the EU in the data in the OECD health database and in New Cronos.

Although averages are often used as a kind of normative goal, they are not a yardstick. They should not be seen as a target standard, and deviations from the average are neither good nor bad. Averages become more solidly based if the spread around them is taken into account. However, many policy analysts use averages without presenting spreads or deviations from the mean. What is much more important, however, is the content of the data making up the averages. Only if this content is the same, will the average and the spread around the average be meaningful.

9.4 Personal health care percentage of GRP¹⁾ and GSP²⁾ in the USA, 1998



¹⁾ GRP: Gross Regional Product.

²⁾ GSP: Gross State Product.

Source: Health Care Financing Review, Summer 2001, table 4.

At face value, without any additional information it is not possible to say that the data provided for the separate states of the USA on the basis of the state health expenditure accounts are more comparable than the data provided for the EU on the basis of the national accounts or to the OECD based on national accounts or on health accounts. In the knowledge that the data in the USA concern only health care expenditure in the USA, and that all the data are collected and estimated by one single institution (the National Health Statistics Group) the differences are attributable to levels in state population (larger populations spend more), age distribution (older people spend more), personal income (higher incomes spend more) and insured status (insured people spend more) and the physical location of the provider: the lack of borders makes it easier to import and export medical care between states in the USA than in the EU.

Although it might be expected that the differences between the states of the USA will be smaller than, say, between the countries of the EU or OECD, the data do not confirm this. Probably the methods and interpretation of methods and classifications may not be as unified as supposed. Apparently, even in the USA, with its high level of co-ordination and concentration of data processing, Kanavos may have been right in concluding that differences "... may be due to methodologies according to which data are collected, compiled and analysed." (Kanavos 1996, page 15).

9.5 *Health care data in the OECD health database*

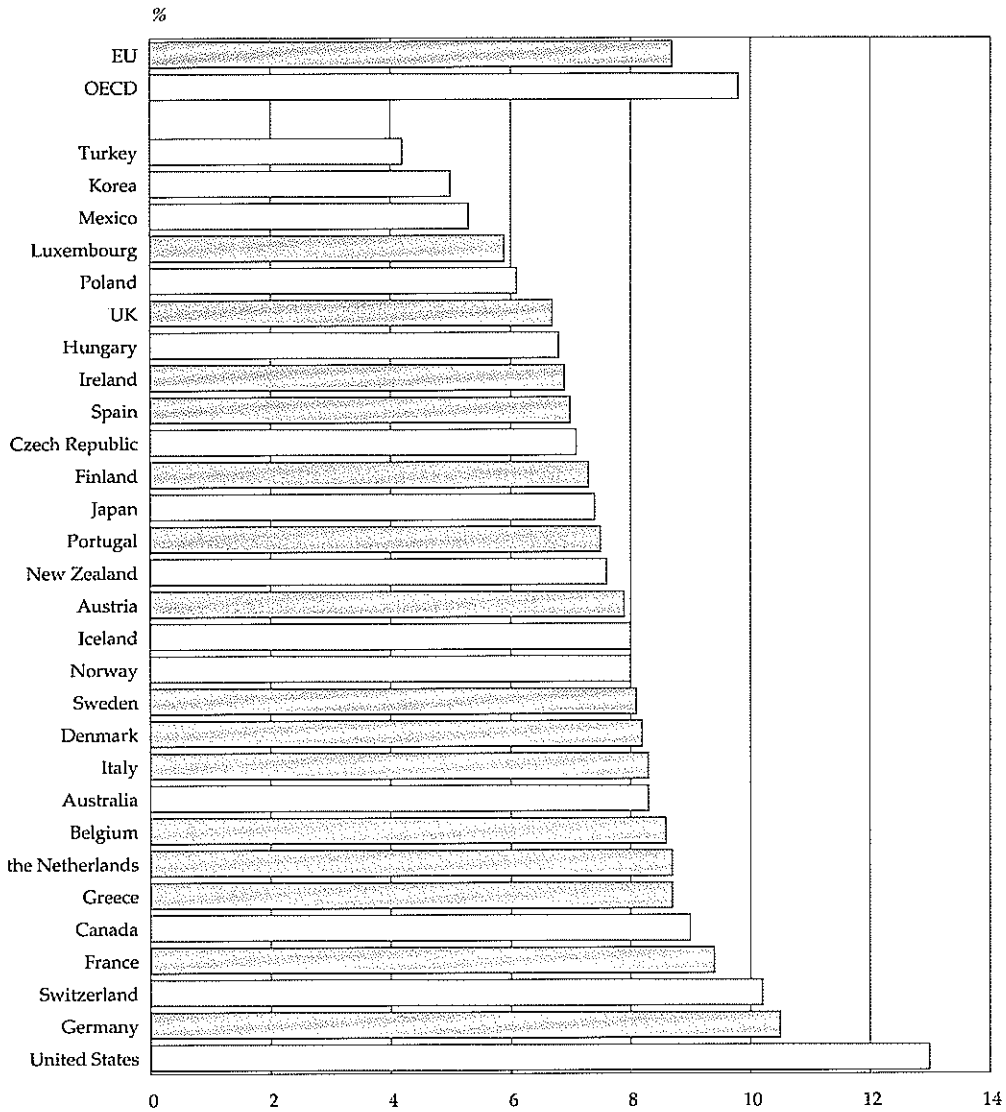
In section 9.2 data from the OECD health database were presented for the EU member states. Of course the OECD health database also contains data for non-EU countries. Some aspects of the OECD health database are presented below.

The OECD health database is a very rich source of data on health in all its components. There are many data on health care expenditure, presented in a way to make them as comparable as possible. Graph 9.5 gives the data on health care expenditure as a percentage of GDP for the OECD and the EU areas, as well as for individual OECD and EU member states. The averages presented are weighted averages in which GDP data (in US \$) are used as weights.

Like the sources presented earlier in this chapter (New Cronos, OECD services: statistics on value added and CMS Data), again it is difficult to say anything about the level of comparability in terms of shares of GDP. It is no more possible to determine the level of comparability of the data between USA (13.0 percent) and the UK (6.7 percent) than to determine the level of comparability between Austria (7.9 percent), Iceland (8.0 percent) and Norway (8.0 percent).

It is obvious in the data as presented in this chapter that "... in the comparative collection of national data it is possible that, based on different boundaries in the

9.5 Health care expenditure of OECD and EU member states, percentage of GDP, 1997



THC: Total Health Care (including investment in medical facilities).

Source: OECD Health Data 2001.

national statistics and/or methods, differences appear in the publications of international organisations ...” (Schneider 1998, Vorwort). What is more, these differences appear in data based on national accounts, the OECD health database (a mix of national accounts methodologies and health accounts methods) or other health accounts methods.

9.6 Time series analysis

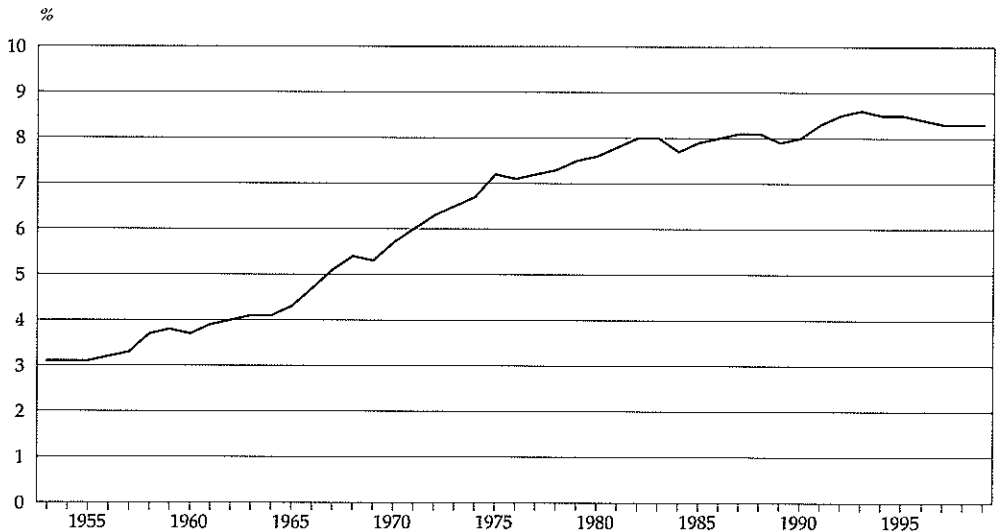
As set out in chapter 1, one important aspect of comparability of health care expenditure data is comparability in time. The goal of time series analysis (of both foreign countries and one's own country) is to learn something about the positive and negative developments in the field of health care and their possible applicability in the own situation. Time series analysis is in the first instance limited to a single country.

Within a single country comparability of the data is highly correlated, or even identical, with the consistency of the area described, and the consistency of concepts and definitions used. Such consistency is not always guaranteed over a longer period of time.

For the Netherlands, for example, a consistent time series of expenditure on health care was created for the period 1972–2000 in the statistics on cost and financing. In the period 1972–1992 no major changes appeared in the area of health care described, methods applied did not change and no major changes were introduced in the statistical data. In the period before 1972 different methods were used in the statistical process and the area described underwent changes. After the beginning of the 1990's changes in the health care area became visible and ultimately led to the creation of new statistics on care accounts.

Looking at the data (expenditure on health care as a percentage of GDP) for the whole period 1953–1998 the reasons explaining the growth in the time series for the

9.6 Health care expenditure percentage of GDP in the Netherlands, 1953–1999



Source: Statistics Netherlands, Cost and Financing of Health Care 2000 and earlier issues, National Accounts 2000 and earlier issues.

Netherlands are only visible at a high level of aggregation. The consequences of increasing prosperity and the expansion of the packages of provisions in health care during the 1960's can be detected in the growing share of health care in GDP. The introduction of the Exceptional Medical Expenses Act (accounting for almost 8 percent of total health care expenditures in 1968), however, can no longer be recognised in the data. In the 1980's and 1990's a levelling of the growth is visible resulting from the economic slowdown (in the 1980's) and the preparation and introduction of the EMU in the 1990's.

The ratio of health care expenditure to GDP is probably not the best way to analyse a country's time series, because both the numerator and the denominator influence the results. The interest in a time series on health care is in developments in health care expenditure, the possibilities to identify large or small results of political measures. Growth in total health care expenditure can be broken down into population growth, price increases and the remaining growth (utilisation changes and performance) ³. Remaining (or real per capita) growth of health care expenditures can be calculated as total growth adjusted for population growth and for inflation. The efforts in this area of research need to be intensified (more in-depth analysis) and expanded to include more explaining variables.

With all the knowledge available on the area of health care, it is hardly possible from this graph to draw any conclusion on the results of policy measures on real growth of health care expenditure. The graph presents a 45-year period, and reflects results

9.7 Real growth per capita of health care expenditure in the Netherlands, 1953–1999



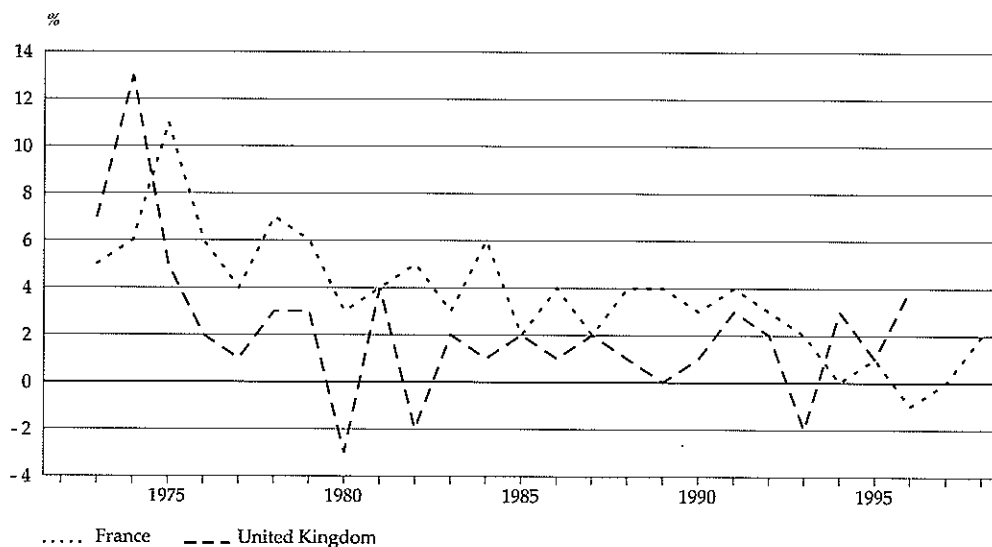
Source: Statistics Netherlands, Cost and Financing of Health Care 2000 and earlier issues, National Accounts 2000 and earlier issues.

of policy measures with a large impact, both in the area of health care and in the economy in general. After a period of expansion in the 1950's and 1960's, the consequences of the oil crises in 1973 and 1976, and the introduction of the budgeting system in the beginning of the 1980's are visible.

Health care policy measures are usually more specifically aimed at changes within the health care sector, for example shifting the use of health care from in-patient to outpatient providers, or shifting financing from public to private sources. The measures have hardly any effect on the level and development of total health care expenditure. If the consequences need to be evaluated at this level in the national situation, much more detailed information is needed. Analysis of such information at cross-country level compounds the level of difficulty extremely.

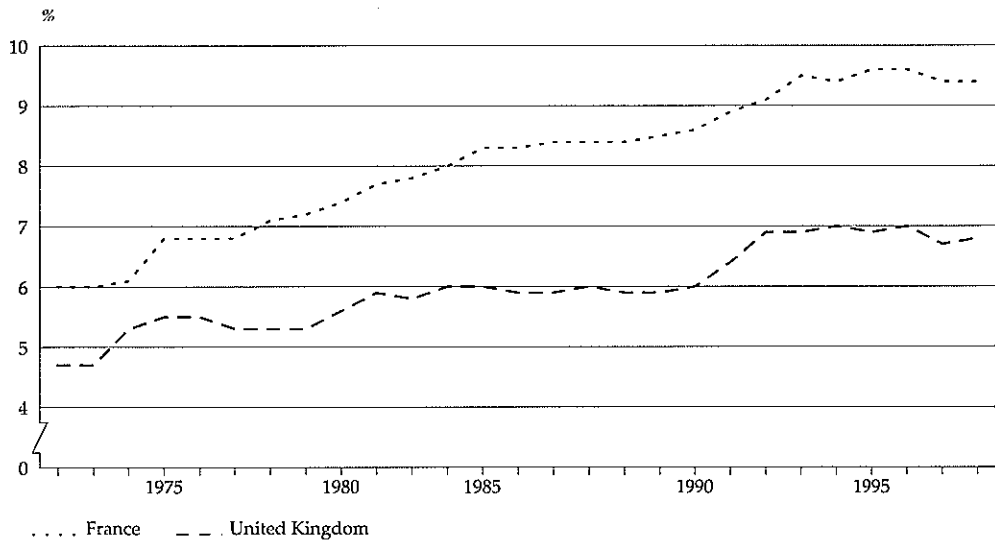
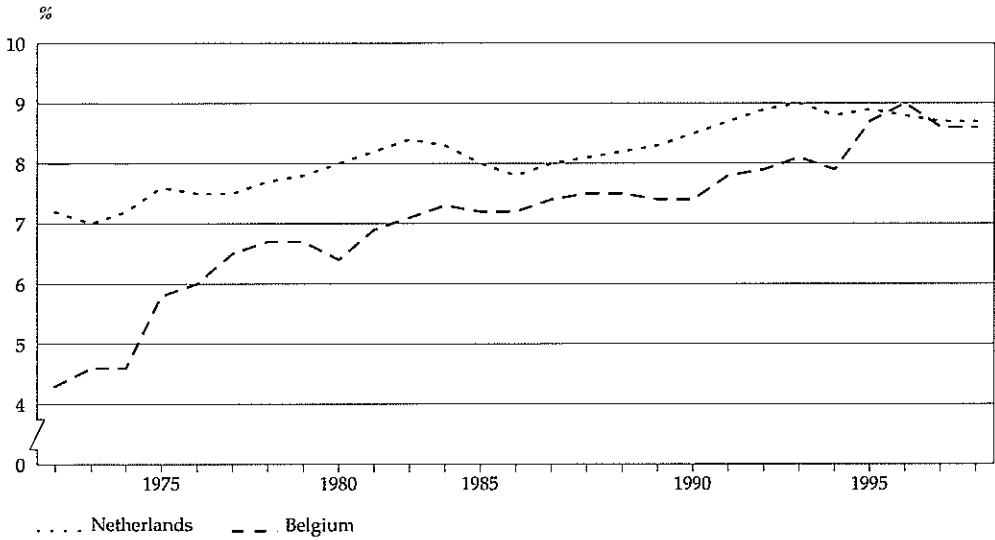
What is clear from the analysis above is that a lot of information and knowledge is needed on the data, the structure and system of health care provisions as well as policy measures, to be able to provide any conclusions on the data for the Netherlands. The multitude of policy measures as well as their simultaneous introduction make it extremely difficult to gain insight into their results. The possibilities for other countries to use the results of the Netherlands in their own situation depends largely on the organisation and financing of health care and non-health care related factors like cultural circumstances and political possibilities. Graph 9.8, with real growth rates for two countries based on data from the OECD health database, illustrates the problems in using these data.

9.8 Real growth per capita of health care expenditure in two OECD countries (France and the United Kingdom), 1972–1998



Source: OECD Health Data 2001.

9.9 Health as a percentage of GDP compared for four EU Member States (the Netherlands and Belgium versus France and the United Kingdom), 1972–1998



Source: OECD Health Data 2001.

Usually in cross-country time-series the data are presented as a percentage of GDP during a period of time. This is not the best way if the goal is to learn and possibly apply policy measures based on these results. Graph 9.9 presents a selection of time series for the period 1972–1998 (for which period data are included in the OECD health database for the Netherlands in a consistent manner).

Looking at the time series for Belgium and the Netherlands, it can be seen that the respective trends converge, leading to a similar share in GDP for both countries in the 1990's. For the UK and France the data seem to suggest that the trend in health care expenditure as a share of GDP is diverging, leading to a gap of around two percentage points between the two countries.

In general terms it can be stated that for using time series, especially those of other countries, a lot of background (or meta) information is necessary to be able to draw the right conclusions. This is compounded by the fact that time series are only relevant for items or clusters of items that contain the same sets of activities, so that what is compared across countries has the same content.

9.7 *Final remarks*

Comparability across countries of data from whatever source is not automatically a given fact; but not many users are sufficiently aware of this. Often the data presented are taken as valid and comparable across countries at face value.

A favourable exception is presented in Kanavos 1999, where it is stated that "Despite the results obtained ... the chapter also highlighted the limits of the available data sources, particularly the OECD databank in actually conducting comparative research for a number of reasons. First [and most importantly for cross-country comparability] because the methodologies applied in the collection of data vary by country" (Kanavos 1999, page 192).

Methodologies applied should be almost identical in the construction of the national accounts. However in comparing the results based on national accounts data it can be noted that:

- if health and social care data, presented as a share of GDP, are comparable, then some strange country results are apparent (e.g. Denmark compared to Italy);
- if, on the other hand, it is concluded that these data are not comparable, then some strange interpretations of methods of existing and agreed frameworks (like SNA/ESA) are used in some countries.

Cross-country comparability of health care data as presented in this chapter is basically undetermined because firstly no standards are available to be used as yardsticks, and secondly the content of the data to be compared is not identical. The data based on national accounts, a source of data that is internationally the most frequently used standard on national aggregates, do not by definition present cross-country comparability. There are two reasons for this: the level of aggregation used, more specifically the way classifications are interpreted in various countries, and the contents of the units classified in terms of activities.

Even using data collected by one single organisation using uniform methods and classifications, for example the National Health Statistics Group in the USA, does not make it possible to determine whether the differences recorded in health expenditure shares between the various states are completely explained by health care system differences.

An analysis of a time series for a country can provide insight into consequences of actions with a large economic impact (in terms of expenditure). However, if it is to be used for conclusions on policy measures within the health care area, the required level of knowledge and detailed information is extremely high. Usability of this information across countries implies the same level of knowledge and information of these foreign health care systems as on one's own country.

It is not clear in advance which source provides the best cross-country comparability: a more detailed source like the OECD health database or a less detailed but more co-ordinated source like national accounts. What is clear, however, is that international comparisons of health care expenditure do not provide a definite answer on:

- the level of expenditure that is appropriate;
- the ideal mix of services constituting health care;
- a normative level of efficiency, efficacy, equity, access or
- any similar questions to be posed in the area of health care expenditure.

What international comparisons of health care expenditure should provide is a definite answer on the expenditure associated with the contents of the units of production and consumption in the health care area across countries.

Without any detailed background information on the content of the data presented, it is completely impossible to determine the level of comparability within a source, between sources and the influence of the human factor in various countries for different sources.

Notes:

- 1) Production value is the value of the goods and services produced measured at market prices (basic prices or production cost) seen from a supply or supplier's point of view (derived from ESA 95, see also chapter 3, note 11).
- 2) Expenditure (derived from ESA 95, page 51) is a way to express the exchange of goods and services, i.e. when the purchaser incurs a liability from the seller, for cash from the paying or purchaser's point of view.
- 3) See for an explanation of the effects on total health expenditures e.g. OECD 1985, *Measuring Health Care*, page 15, OECD 1990, *Overview*, pages 12 and 13, and *Cost Containment in Europe*, pages 32 and 33).



10. Existing frameworks: conclusions

10.0 Questions

The first part of this study deals with the fact that expenditure as presented in various existing databases is not or not completely comparable between countries. For existing international frameworks this problem can be translated into the following set of questions:

Is comparability an item in existing frameworks, in other words is comparability across countries an issue in the description and construction of existing international frameworks? If not, or not ideally, has comparability improved in newer versions?

Are the starting points and criteria mentioned in these frameworks favourable or detrimental to comparability in general and comparability on health care in particular?

Does the use of more detailed systems like satellite accounts or health accounts increase the level of comparability?

10.1 Introduction

In the previous chapters existing frameworks, which are used in international economics and can be used in the analysis of health care studies, have been presented and discussed. Existing systems in this context are frameworks developed since the 1950's and broadly applied in many countries. In this chapter the conclusions on the use of these frameworks for the problem of comparability are presented: general similarities and differences with regard to comparability; problems regarding consumption and production; and lastly, the results of the previous chapters.

10.2 Usability and comparability

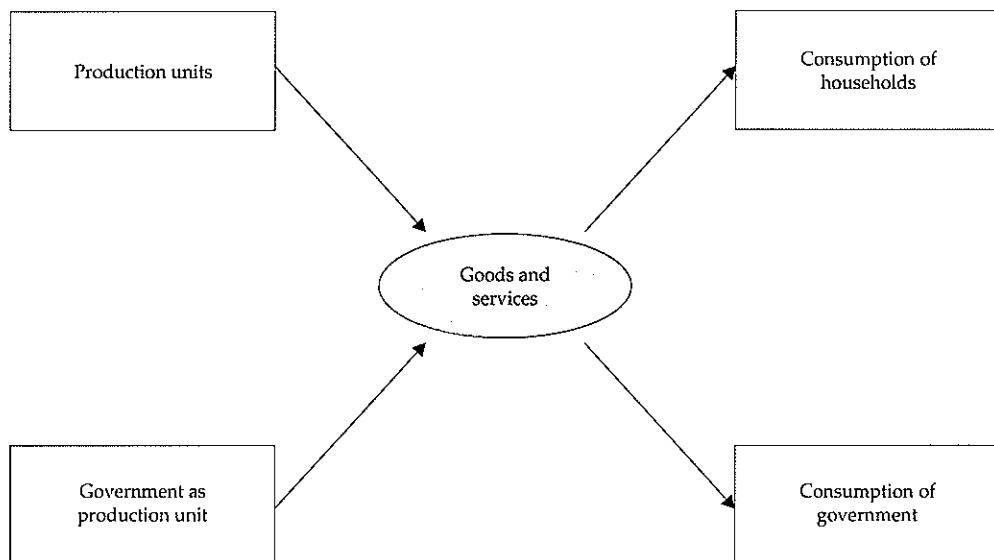
The usability of the frameworks is limited to the health care branch and the basic object of this study: production and the related items production value and expenditure. In disaggregated systems like satellite accounts and health accounts, one of the main goals is to indicate what it costs society to maintain a health care

system (see section 10.2.2). Just as important is how units are classified (see section 10.2.3) and the basic classification used (see section 10.2.4). A last point related to incomparability is the purposes for which the various frameworks were created and are used (see section 10.2.1).

10.2.1 Frameworks and relations

All frameworks described have to do with economic relations, more specifically the relations between producing units and consuming units as schematically presented in figure 10.1.

10.1 Relations between production and consumption



As stated in chapter 3 on the System of National Accounts (SNA) and the European System of Accounts (ESA) the macro-economic models deal primarily with relations between independent economic actors at the highest possible national level of aggregation. In SNA and ESA the relations between institutional sectors are of prime importance. Health care is not a sector in SNA/ESA, but a production group in the ISIC/NACE (International Standard Industrial Classification of All Economic Activities/Statistical Classification of Economic Activities in the European Community) production classification.

Satellite accounts (chapter 4) do not refer to macro-economic relations between institutional sectors, but present a more detailed picture of certain (usually social) phenomena that may be part of all institutional units and all institutional sectors as

defined in SNA. But as in SNA the relations between the phenomena described and the rest of the economy are very important. For a satellite account on health care, the production value of health care goods and services must be retraceable in SNA.

Like satellite accounts, health accounts (chapter 6) describe a certain set of phenomena (constituting health care), presented in the macro-economy. Health accounts present all activities concerning this phenomenon and the rest of the economy, but without the constraints of the definitions, classifications and valuation peculiarities used in SNA. For health accounts the relations with other actors in the economy are also important; these may be, but are not necessarily, institutional sectors.

Lastly, the OECD health database (chapter 7) tries to present an integrated consistent and comparable picture of the health branches in all member countries. Of prime importance for the OECD (Organisation for Economic Co-operation and Development) health policy unit is comparability of health care and health care expenditure across countries. For all other frameworks (SNA, ESA, satellite accounts or health accounts) this is not the case. Health accounts try to present as complete a picture as possible of the branch within one country; satellite accounts try to do the same using the framework of SNA, which in turn contains comparable data, but focuses on macro-economic aggregates and not on comparability across countries for a certain class or group of ISIC and NACE.

10.2.2 Frameworks and valuation

The way transactions are valued and the way the units (especially units of production) are determined and classified, are crucial to the differences between the various frameworks.

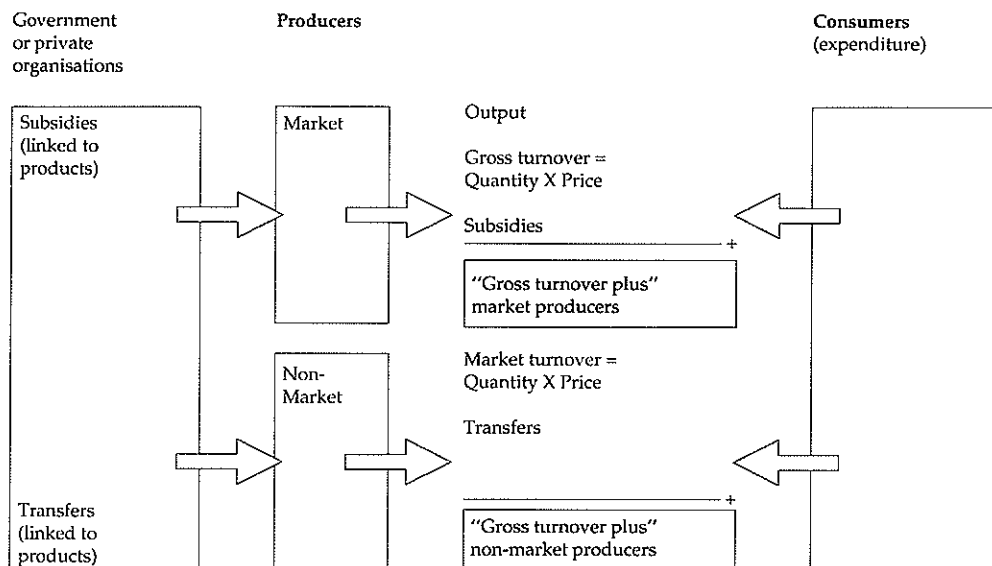
A basic term used in the main system (SNA/ESA) is production value, i.e. the representation of the value of the "... goods and services produced for sale on the market at economically significant prices." (SNA 1993, page 152). Valuation should be at basic prices (excluding deductible indirect sales taxes like Value Added Tax, VAT).

In satellite accounts as well as health accounts production value is usually interpreted as the expenditure by society (consumers, government, NPISH and others) on the health care branch. In satellite accounts, however, this expenditure notion must be recalculated to production value in SNA terminology to make a link possible to the general framework.

The confrontation of this description of expenditure for society with the definition of expenditure in SNA terminology gives some interesting results. In SNA 1993 expenditure is defined as "... the values of the amounts that buyers pay, or agree to pay, to sellers in exchange for goods or services that sellers provide to them or to other institutional units designed by buyers." (SNA 1993, page 207). This definition

of expenditure is only in agreement with the general notion of expenditure used if transfers or subsidies from government bodies or other institutions can be seen as an exchange of goods and services between the units supplying the money and the units receiving the money (reflecting the gross turnover of these units). These financial flows are part of the business receipts but are not always reflected in the 'production value' of goods and services (see figure 10.2).

10.2 Valuation and gross turnover

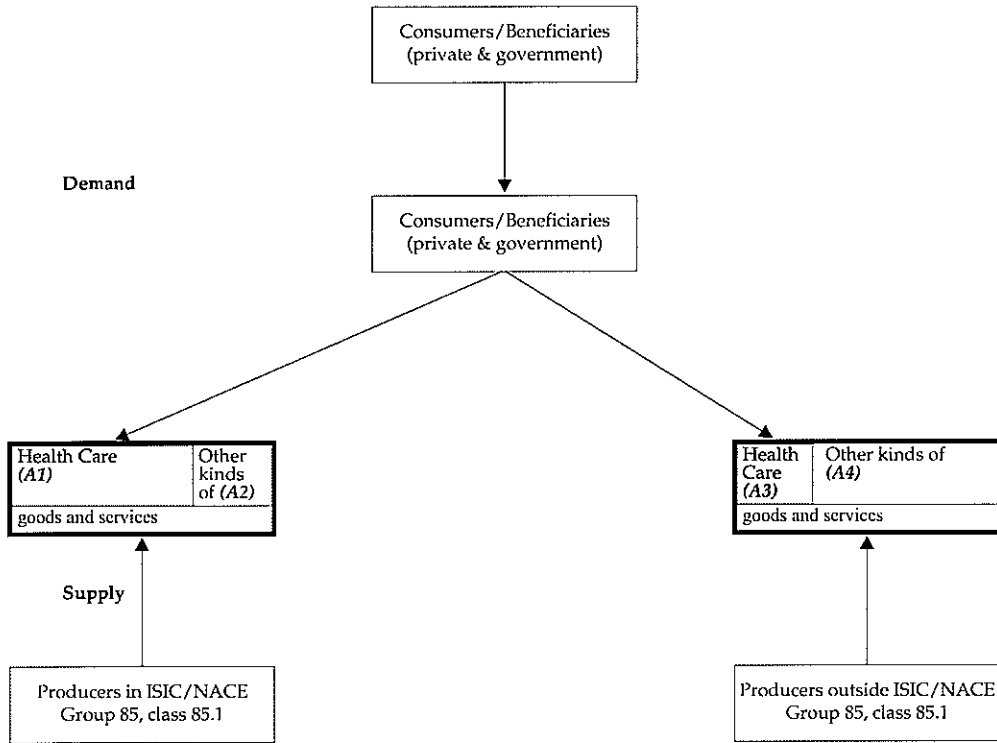


A minor point of discussion is created by third party payers. In SNA "... expenditures are attributed to the units which ultimately bear the costs as distinct from the units that may make the payments." (SNA 1993, page 207). SNA is not interested in the primary payments arrangements (initial payments made to providers and other actors in the system). Certainly in health accounts (and to a lesser extent also in satellite accounts) both the ultimate payers and the primary payers are important. To make sure that all expenditures are accounted for the final link between the paying units and the suppliers of health care (including the suppliers of management and control) is indispensable.

10.2.3 Frameworks and units

The last, and perhaps most important aspect for differences between the frameworks concerning the presentation of health care is how the units are used. Figure 10.3 gives a schematic presentation of these differences.

10.3 SNA, Satellite Accounts and Health Accounts: production compared



- A1: Health care goods and services produced by units with a main activity Health Care
- A2: Other kinds of goods and services produced as a secondary or ancillary activity by units with a main activity Health Care
- A3: Health care goods and services produced as a secondary or ancillary activity by units with a main activity outside the area of Health Care
- A4: Other kinds of goods and services produced as a main activity by units with a main activity outside the area of Health Care

In SNA and ESA units are created using the main activity principle. This means that based on the activities performed by any institutional unit, the most important activity (in terms of output share, production process, use of inputs or any combination of these criteria) determines the classification of that unit in the ISIC or NACE classification. As a consequence of this main activity principle, secondary activities outside the health care branch are included in the production value of the units classified in ISIC 85.1 (health care). At the same time, however, this principle excludes all health care activities performed by units outside the health care branch. Looking at health care in SNA terminology, the production (value) would be the value of the goods and services presented by area (A1) and (A2) in figure 10.2. This would be the equivalent to the production value of ISIC/NACE group 85.1. For satellite accounts this main activity principle is adjusted to meet the needs of the users better. For all units classified in the health care branch (ISIC 85.1) secondary

activities outside health care are still included in production value. To create a more complete picture health care activities of all units outside the health care branch, either performed as a secondary or as an auxiliary service, are included.

In satellite accounts production (value) should equal by and large the areas (A1) and (A2) in figure 10.3 as well as area (A3). The latter area should represent all health care activities performed as secondary or ancillary services.

For health accounts a different area again would be designated as health care production, namely areas (A1) and (A3): health care goods and services provided by all producers in the economy, irrespective of their classification in SNA/ESA terminology. The place in the classification of ISIC is not important in health accounts. All activities in the health care field ought to be included in the accounts, whether they are performed as a main, secondary or ancillary activity.

A last remark to be made concerns the contents of the area health care goods and services performed by producers outside ISIC/NACE 85.1. Although in both satellite accounts and health accounts this area is part of the production of health care, this is not to say that the contents of this area are the same for both approaches. Boundaries may differ. For both approaches, for example, it is possible to include or exclude units performing activities in the area of health-related research and development, health care education and training, environmental health care activities, and other health care related activities.

10.2.4 Frameworks and classifications

The classifications used in the various frameworks constitute another important difference. Classifications have a major influence on the boundaries of the economic relations described. ISIC and NACE are the main classifications used in the SNA and ESA to categorise units of production. This means that retailers of medical goods are not included in the health care branch. The same is true for government as a producer of health care and for all publicly and privately funded health care insurance companies. Of course not all management, control and administration of health care is included in the health care class, only those instances that are an integral part of health care providers.

As in every day life the boundary of the health care branch is broader than the one used in SNA. Both satellite accounts and health accounts use a wider boundary, usually including retail trade of medical goods, government health care production and insurance in health care. Financial information on the topic of administration and management of health care activities is usually included.

Another boundary problem creating differences between the frameworks is the way household production is treated. SNA/ESA excludes household production for own use, with some exceptions. For high per capita income economies like those in the EU, the way household production is treated is relatively unimportant. However, for comparison with countries outside the EU with lesser developed

economies, registered differences may be caused by the way in which household production is treated. Many activities in the health care area performed by professionals in professional settings in developed economies may well be performed informally by households in developing economies.

However, in some countries, like the Netherlands, government is encouraging households to solve (within legal boundaries) medical (care) problems in a way they see fit. In the near future this could mean that a substantial part of the total health care expenditure circulates between consumer households. For the purposes of comparability the economic boundary of the health care system will require some adjustment in that case.

Classifications by purpose (functional classifications) are used in the framework of SNA/ESA for the description of consumption. Adding collective consumption (presented in the Classification of Functions of Government, COFOG) to individual consumption (presented in the Classification of Individual Consumption by Purpose, COICOP and in the Classification of Purposes of Non-Profit Institutions Serving Households, COPNI) gives total consumption (final and expenditure) for national accounts. Consumption defined in this way does not have to be identical to that used in health accounts.

10.3 Frameworks and the differences in production (value) and consumption (expenditure)

Basic to the description of the usability in cross-country comparisons of health care are the terms production and consumption. Even not taking into account the possible differences following the boundaries of the area described (research and development, education, etc.), the differences between the approaches regarding these two items are important.

In SNA terms production is valued at basic prices, excluding VAT and similar production taxes. In the health care branch, wage subsidies are excluded, as is all financial intermediation (attributed to the financial sector at national level). Identity of production and consumption inside SNA/ESA is not guaranteed in the area of health care, because the production of health care goods and services is limited to production units in ISIC 85.1 and consumption of health care goods and services also includes secondarily produced health care items.

Because the area described in satellite accounts differs from that in SNA/ESA, the production in these two approaches also differs. Consumption of health care measured in goods and services needs to be adapted to the boundaries described in these two approaches.

Including the retail trade in medical goods creates a problem with the connection to the general framework of SNA/ESA, where retail trade is valued at its 'trade

margin'. Providing a picture of the value of production of retailers as the final link in the production column, the trade margin is not what the production column medical goods provides to its customers. Sales at market prices would be a much better indicator. If the trade margin alone is included in satellite accounts this would also create a difference with the consumption of medical goods valued at 'market prices'.

In health accounts production is usually valued at 'market prices' and, as there is no market in many countries, the value is estimated as the total of the receipts of the units of production in the health care area. These receipts include VAT (if applicable), subsidies, interest receipts, etc.

Actual consumption and consumption expenditure are among the most difficult topics in the framework of SNA/ESA (see chapter 3, section 3.6.3 consumption, expenditures and financing). Using COFOG and COICOP in all countries to calculate total consumption does not automatically give comparable results, as countries apply the rules and criteria differently: using units by main activity as classifying principle, parts of units if appropriate and possible, and activities or services provided.

In health accounts consumption expenditure is distinguished from financing. Consumption is related to the individual and collective units that benefit, financing relates to the units that initially and ultimately pay for the production.

Consumption and production differ in the various frameworks because of differences in the area described, but also in the way these items are measured and valued.

10.4 Conclusions based on results of the frameworks

10.4.1 Conclusions on SNA and ESA

In the absence of a theory of comparison or comparability, a practical way to investigate the possibilities of comparability across countries is by using existing international frameworks. However, the use of internationally agreed frameworks in itself does not guarantee comparability across countries or in time. Using frameworks implies a multitude of decisions on, for example, compiling and calculating methods to be used and on boundaries of the area under study.

National accounts, although used in international comparisons by all international organisations and almost all countries in the developed western world, were not created for the sake of cross-country comparisons. The purpose of SNA is to create complete and consistent national income and product accounts, using agreed conceptual standards and classifications. Cross-country comparability is limited to

macro-economic aggregates. In the last 50 years, and especially in the last revision, the opportunities for comparability have been considerably improved at this level, with increased harmonisation in definitions, concepts and classifications.

However, comparability is one of the main goals of national accounts in the EU area in conformity with ESA, because at the European level the results of national accounts are used for European policy on the national financial contributions and the re-division of funds. Comparability is implicitly present, by the application of strict rules and regulations, in the construction of national accounts. In ESA comparability is strengthened by the explicit use of basic prices in the valuation of transactions, but is weakened by the introduction of thresholds. Especially in areas with large numbers of small units below the thresholds, the results may be underestimated. In non-market services and more specifically in health care large amounts of small enterprises are present in some countries and less in others.

National accounts ensure a high level of comparability at the macro-level; comparability at a lower level of aggregation is not a goal. The lowest level in SNA and ESA for which comparable data ought to be presented, is the level of the institutional sectors. At industry group level the availability of data can be problematic, as well as the contents of the groups compared across countries. Contents of industry groups across countries are by definition not the same and therefore difficult to compare.

SNA and ESA are important because of the agreed use of concepts, definitions, terminology, units and classifications. However, classifications, units etc., do not create comparability. The use of these instruments only makes the differences between countries visible and thus opens the possibilities to achieve greater comparability of statistical results. To realise greater comparability across countries, the variables to be compared need to be alike in terms of activities or clusters of activities performed.

10.4.2 *Conclusions on satellite accounts*

By definition, satellite accounts offer more possibilities for comparability than SNA and ESA, as they can be used to investigate topics of social concern. As these topics are often associated with non-market production and government influence or control, a lower level of aggregation is possible, as well as more detail.

Satellite accounts also offer – in theory – more flexibility in classifications, units and concepts. As already mentioned, the use of classifications and concepts is not identical to comparability, but the use of uniform classifications is favourable for comparisons. More flexibility does, however, foster usability, which in turn can show the differences across countries and thus provide possibilities for increased comparability. Opposite this favourable condition, it must be noted that flexibility in itself offers possibilities for deviating concepts and classifications, and thus larger differences between countries and less comparability.

Satellite accounts were developed as early as the 1970's, but are still only regularly used by a limited number of countries. A few important reasons for this are that constructing satellite accounts is very laborious, and only worthwhile if there is a clear added value for the users. The results of satellite accounting are probably not very usable for managers and policy makers in the health care branch, because the information supplied at the macro-level is not the information needed at the meso- or micro-level. Apparently satellite accounts were not developed to meet the (undefined) needs of the users, but for the convenience of the producers of the data. For cross-country comparability, satellite accounts do not offer the desired solutions. Comparability across countries on areas of social concern, like health care, is not easily possible within the framework offered (having to define characteristic activities, characteristic products, connected goods, etc.). The largest problem encountered is that the levels of entities compared are only very mildly comparable because the activities performed by those entities differ between countries.

10.4.3 *Conclusions on health accounts*

Health accounts, also sometimes called national health accounts or functional health accounts as opposed to the more institutional accounts produced within the framework of SNA, offer by definition very small opportunities for cross-country comparability. The reason is that they describe an area that is largely determined by national features, national sentiments and national history. Especially in this area it is obvious that data are collected for specific national purposes, not for international comparability. Only by using special research techniques in a small case experiment a high level of cross-country comparability can be achieved (see e.g. the studies of Abel-Smith in the 1960's).

10.4.4 *Conclusions on the OECD health database*

The OECD health database includes data of all member countries. These are statistics based on national accounts, satellite accounts and numerous different systems of health accounts. Although the secretariat of the OECD does its best to create comparability between these various results, the number of footnotes and other metadata provided show clearly that comparability, even in the area of health care expenditure, has not yet been achieved. There is a distinct need for more harmonisation, co-ordination and standardisation in the health care area in order to get better comparability.

One of the most important conclusions of the discussion on the OECD health database concerns the contents of the units described. Units with exactly the same names in various countries do not automatically mean that these units perform the same sets of activities and hence are comparable in their expenditure.

10.4.5 *Conclusions on functional classifications and the statistical process*

Functional classifications in themselves can be a solution for greater cross-country comparability for consumption expenditure, as they all incorporate health care and

expenditure. The greatest disadvantage, however, is the level of flexibility and freedom offered by the classifications. Freedom to use different methodologies to construct and use a classification does not increase cross-country comparability. Combining various classifications that can be constructed by various methods will already present problems within one country. Comparing two or more countries does not result in comparability of expenditure.

The three constituent parts of the statistical process are concerned with data availability (input), user demands (output) and data processing (throughput). Only the throughput process can ultimately be controlled by the statistical authority. Cross-country comparability is not a primary concern (either politically or statistically) unless it falls under a European regulation, which health and social care is not. Added to the fact that statistical data are primarily collected for national administrative purposes, this means that cross-country comparability is still under pressure.

10.4.6 Conclusions on data collections

Data on health and social care as presented by various sources (like Eurostat or OECD) are not automatically comparable, not even if only sources based on SNA/ESA are used for comparisons. SNA/ESA dictated methods, classifications and procedures, do not always lead to uniform results (partly because the human factor is involved in interpreting rules and regulations). Most researchers do not always take into account that data used in their models are not as comparable as they would like.

In comparisons of health care data, whether on production value or expenditure, no agreed content of the terminology is available and the boundaries of the area under study are not internationally agreed upon. It is impossible to say beforehand, without any in-depth research of data and metadata, which source provides the best comparability and what the level of comparability of the data actually is.

The construction of a national health expenditure account is still very complicated, even with the use of modern high-tech IT solutions. Information technology presents possibilities to use different definitions and classifications nationally and to provide the data needed for these different points of view. On the other hand, there is usually not enough time to perform all these possible exercises. However, rearrangements of national data performed by international organisations are nowadays less acceptable and less accepted, because of the available national level of detail and the technical possibilities.

International use of national data frequently requires a reorganisation of national data to comply with internationally used definitions and classifications. These rearrangements of nationally supplied information used to be accepted. However, possible technical solutions in this area have rendered these rearrangements by international organisations less acceptable for national authorities.

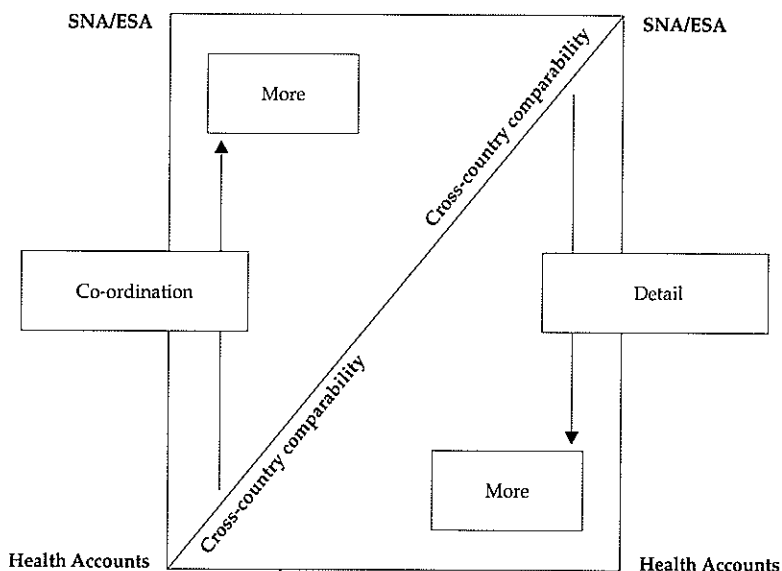
10.5 Final remarks

Expenditure as presented in the various frameworks on the topic of health care are not comparable across countries. There are problems with the boundaries of the area described, classifications used to determine the items and variables to be included and described, and definitions. All these differences can be traced back to different users and different user demands.

Cross-country comparability is an issue within existing frameworks, although the level of aggregation is not in accordance with the level requested by the potential users of the results. Furthermore the classifications and definitions used are not always compatible with the needs of users either. This means that comparability is a relative concept, determined by the object to be compared and the purpose of the comparison.

In the newer versions, the possibilities for comparability across countries have been improved, not only through better classifications and definitions but mostly by more voluntary co-ordination, co-operation and understanding of each other's problems and possibilities. The OECD health database in particular has profited from co-ordination and co-operation on a voluntary basis. It is also clear after 25 years of maintaining the health database that the possibilities to improve comparability are limited and new ways need to be explored.

10.4 Co-ordination, detail and comparability



Comparing the various frameworks discussed in the previous chapters, it is obvious that using more detailed systems presents better opportunities for improved comparability across countries. This means that for health care expenditure, SNA provides least possibilities for cross-country comparability, satellite accounts provide better opportunities, but health accounts, being the most detailed, provide the best opportunities. Opposite the level of detail is the level of agreement on classifications, boundaries, definitions, etc. The less agreement there is in the use of co-ordinated and agreed classifications, boundaries and definitions, the fewer the opportunities for cross-country comparisons will be (see figure 10.4).

For health accounting the limiting factor is the lack of sufficient co-ordination and co-operation on agreed frameworks, boundaries and definitions in the area of health accounting to be used for cross-country comparisons. To realise this goal, more co-ordination and co-operation is necessary between all parties concerned.

Evaluating the conclusions drawn in this chapter on the usability of existing frameworks in cross-country comparisons, the problems to be solved are the same as those presented 20 years ago. In the beginning of the 1980's the problems in health accounting comparisons included among other things boundaries of the system of health care, links to the SNA, health care system and the contents of accounts and the availability of statistical sources (see Levy 1982, page VI, VII).

The two most remarkable conclusions to be drawn from the discussions of the existing frameworks and their possibilities for cross-country comparisons are:

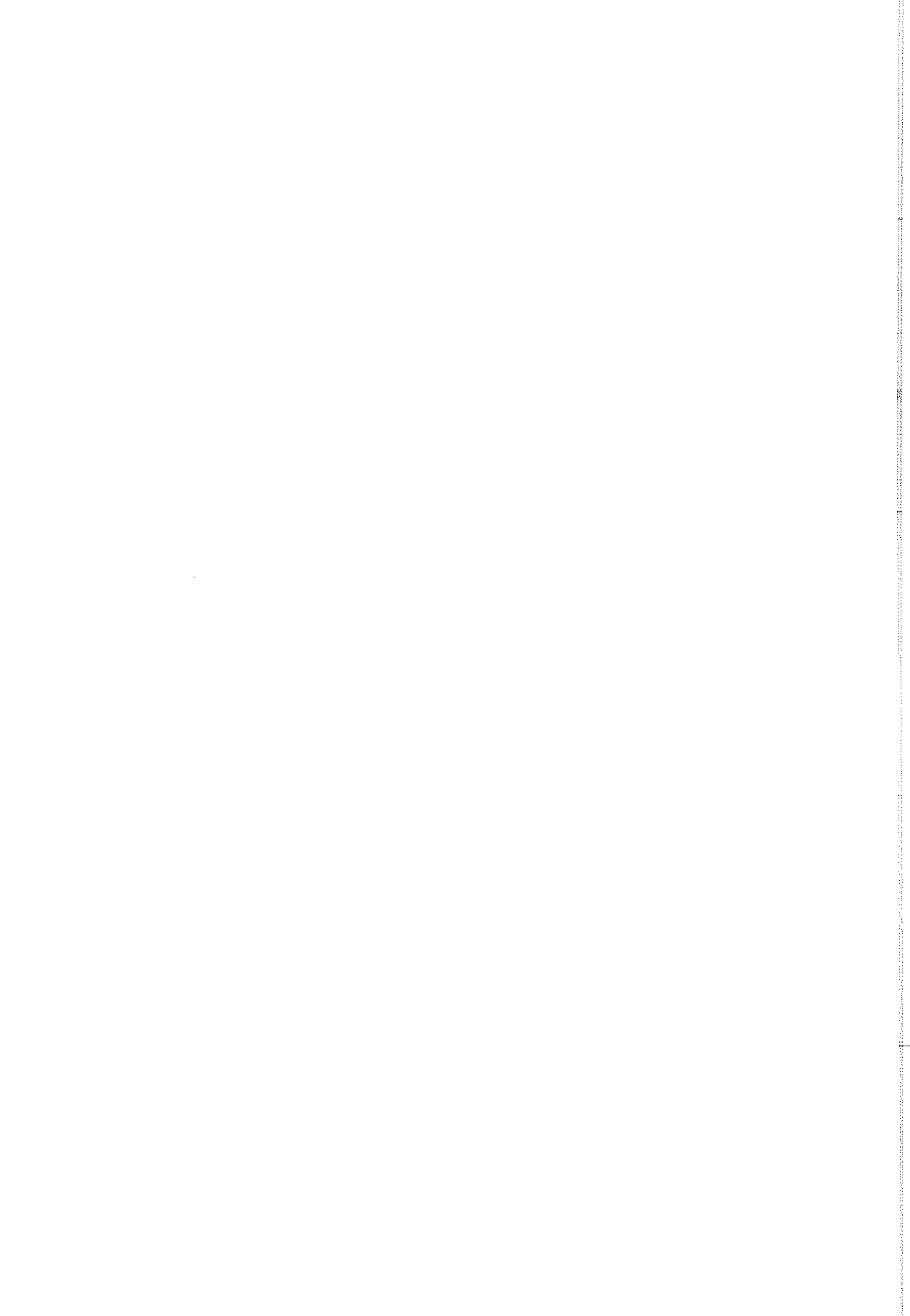
- comparability is a relative concept: determining factors are the object to be compared and the goal pursued with the comparison. Every goal demands its own framework of comparison;
- existing frameworks do not offer sufficient possibilities for comparison generally demanded by users of health care data. Existing frameworks like national accounts offer comparability at a high level of aggregation, next to consistency and coherence in the processing of the data. At a lower level of aggregation (e.g. at the branch level), both goals deviate too much from existing frameworks and objects contain a non-comparable content. Activities are necessary at this lower level of aggregation to make comparisons possible.

It is obvious that for more cross-country comparability on the topic of health care expenditure, new approaches need to be developed. These new approaches need to address the level of detail of, for example, the activities performed by the units under investigation. Only if economic agents or actors with more or less identical clusters of activities are constructed, the results based on these clusters of activities can be made comparable.

The second part of this study investigates the possibilities to enlarge the level of cross-country comparability of health care expenditure data. The most promising

attempts in this area are the method developed by Statistics Netherlands in its Common Comparable Package Project and the method developed by the OECD in its System of Health Accounts. Although the methods differ, both use the starting point of activities as the decisive factor in the construction of comparable data.

*Part 2: Innovations concerning
cross-country comparability*



11. Common Comparable Package (CCP)

11.0 Questions

Does the approach of the Common Comparable Package (CCP) contribute to the improvement of cross-country comparability of health care expenditure?

What are the advantages and disadvantages of this methodology?

11.1 Introduction

The discussions on existing frameworks in the first part of this study (chapters 3 to 9) have made it clear that these international frameworks do not offer many possibilities for comparability of data on health care. The main reason for this is that the objectives of these general frameworks are not in accordance with the more specific goals of health care managers, or politicians. Even if these goals were the same, the objects to be compared do not contain the same contents.

Comparisons take place at the micro-level: for example, an appendectomy in the Netherlands compared with the same procedure in Spain. Surgical nomenclatures or presented Defined Daily Dosages appear in today's state of the art knowledge to monitor activity within and across countries. At the macro-level the major problems are the boundaries. By default the national accounts opt for hybrid concepts that hamper comparability across countries. Two distinct attempts, the Comparable Common Package (CCP) and the System of Health Accounts, address this problem. Both require meso-comparability, obtained through building blocks or through classifications, for which these methods use – as customised as possible to their framework – lists of actors, transactions, products and activities.

In the project on the Common Comparable Package¹⁾ (CCP) the emphasis was put on the contents of the objects to be compared. The CCP project was carried out by Statistics Netherlands at the special request of the Ministry of Health, Welfare and Sports, in the period 1994–1997, with the aim "... to improve the comparability of a package of selected variables." (Mosseveld 1999, page 19). The approach took as its starting point the universal feature included in all frameworks, which is the activity²⁾ performed by economic entities in the economic process.

11.2 General approach

The project promoted a pragmatic approach to comparing the functioning and financing of the health care systems of the participating countries, leaving national definitions and methods of data collection intact. The boundaries in the national systems were not changed.

The description of health care as used in the statistics on cost and financing of health care of the Netherlands was taken as a point of entry in the project:

“Health care concerns the supply of goods and services in the area of medical, paramedical and nursing care:

- provided for human beings suffering from diseases, physical or mental disabilities or limitations;
- related to prevention, diagnostics, treatment and nursing/caring;
- provided by trained experts and/or companies (or parts of companies) set up for this purpose.” (Mosseveld 1999, page 26).

The description of health care relates to health care in a narrow sense as only the providers of health care are included. These providers comprise both private and public health care activities. Health care as presented in the cost and financing statistics relates to a broader description of health care, resulting in a description of the health care system, which includes the activities of administration and management, and thus also expenditure by government and public and private insurance companies on policy, administration and management.

This description of the health care system was taken as an entry point in the comparison. However, it was known beforehand that the national history and national cultural and societal influences of each country meant their health care systems differed with respect to content, organisation and funding.

Because of the different contents of health care systems – more specifically the differences in contents of units providing health care in the various countries – the activities of the providers of health care were taken as a starting point.

As it is not the activities but the trained experts and companies as mentioned in the description of health care that supply information, the logical steps³⁾ to be taken in the data collection process are:

ACTIVITIES ⇒ PROVIDERS ⇒ EXPENDITURE

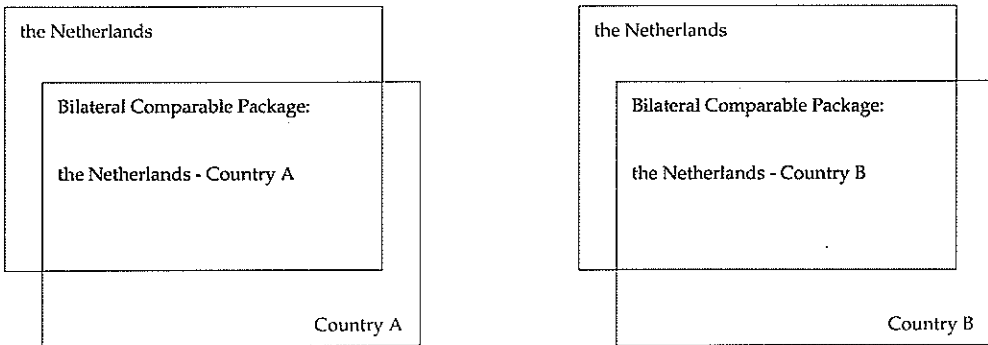
Source: Netherlands Official Statistics, Winter 1998, page 14.

Based on the returned questionnaires (see section 11.3) and the country profiles provided by the participants, the process (see figure 11.1) of determining the contents of the health care systems in terms of expenditure was started.

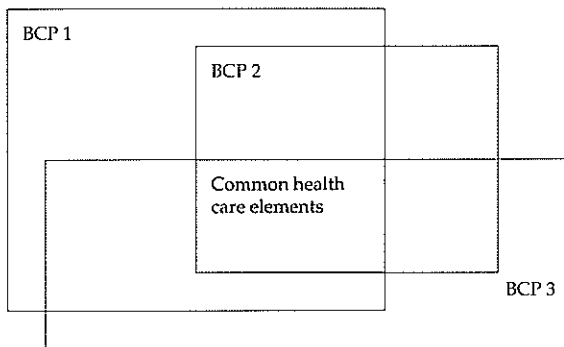
As a first step so-called Bilateral Comparable Packages (BCP's) were constructed, in which the Dutch health care system was used as the key country. The next step

11.1 Process of the construction of bilateral comparable packages and the common comparable package

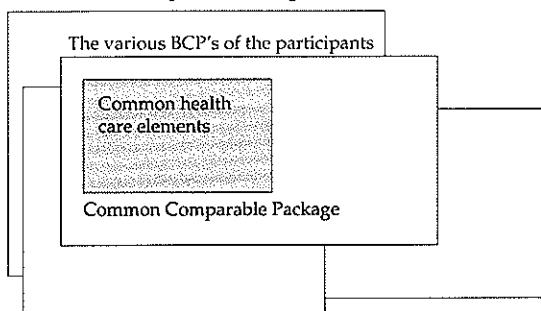
Bilateral Comparison



Towards a Common Comparable Package



The Common Comparable Package



Source: Derived from Mosseveld 1999, page 27, 28.

examined the possibilities for creating a Common Comparable Package (CCP) of health care from the five resulting BCP's. By adding, subtracting and reshuffling provisions, the CCP of health care for the participating countries was constructed. It was more than the summation of the common elements present in the BCP's of the participating countries. The CCP can be used as a reference, a common instrument or concept for international comparisons.

11.3 *Process description and the use of metadata*

The preliminary process of construction of the BCP's and the CCP of health care consisted of the following five steps:

- first of all for all the countries participating in the project so-called *country profiles* were constructed, containing information on the organisation of health care facilities, health care administration arrangements, financing of health care provisions and policy developments.

The country profiles provided among other things a quick introduction to the organisation and administration of the health care system of each country, an insight into the boundaries of the health care system and a preliminary determination of the functions performed in the system;

- secondly, a list of *operational definitions* was considered crucial. However, a complete inventory and description of all definitions used in all participating countries for all health care providers and functions soon proved to be too time consuming. As an intermediate step, an operational list of definitions, based on the descriptions used by Statistics Netherlands (and improved where necessary during the process) was drawn up to help participants understand the nature of the Dutch provisions;
- next, based on a survey of the literature and the information provided by the country profiles *a list of activities* performed in health care was constructed, to provide insight into which activities were performed by which providers in the participating countries;
- respondents returned the list of activities, supplemented with providers of health care who performed them in their national health care systems. These provider lists served as the entry for the construction of *a list of providers of health care (including administration)* as available in every country;
- lastly *matrices* were created. The most important one dealt with the existence of providers in the participating countries, their inclusion in the health care branch or other branches, and the amount of expenditure related to these providers. The information received formed the basis for the determination of the boundaries of care. A second matrix consisted of providers and the financial sources.

A basic prerequisite in any cross-country comparison project is the availability of metadata: i.e. data or information on data. In the CCP approach the country profiles,

the provider descriptions (of the Netherlands) and the activity lists served as primary sources of metadata. This information was supplemented by the numerous annotations supplied by the participants on the matrices (especially providers by expenditures).

11.4 Methods and instruments

During the consultations of the network of experts held in the course of the project, various methods and instruments in the construction of comparability and comparable packages were discussed.

Basically, the method of direct comparison of two health care systems at a time was preferred; direct comparison in terms of observing the contents of both health care packages or parts thereof. At the national level it is important to have determined the boundaries of the health care system and to have listed all the elements included within these boundaries.

Direct comparison resulted in fairly homogeneous packages of goods and services (identical elements in both systems) and differences in other elements. For the differences the consequences in, for example, financial terms were estimated, and ultimately related amounts of expenditure were added to or subtracted from the health care provisions under consideration.

Where no data on elements are available, calculation methods have to be applied. Where there are stable relationships between elements of health care systems, keys can be created and used in the calculations. Keys are "... relative fixed value indicators, expressing a more or less constant relationship between part and total." (Mosseveld 1999, page 36). In the absence of stable relationships between elements, estimates have to be used as an instrument to create the necessary information.

11.4.1 Classifications

11.4.1.1 'Classification of services' applied in the CCP approach

In the course of the project several attempts were made to develop useful classifications of health care services. Ultimately the following 'classification of services' was used to present the results:

- Block I: In-patient, stationary or intramural health care
- Block II: outpatient, ambulatory or extramural health care
- Block III: Preventive health care
- Block IV: Medical goods
 - Block IV a: Pharmaceutical products
 - Block IV b: Therapeutic appliances

Block V: Other services

Block V a: Research and development in health care

Block V b: Education and training in health care

Block V c: Administration in health care

The term block was introduced because the terms used in the classification are partly typically Dutch (intramural, extramural). However, the terms, 'stationary' and 'in-patient' (and 'outpatient' and 'ambulatory' respectively) do not cover the contents as presented in the various blocks completely, although the contents of these expression pairs are more or less the same.

The classification of clusters of health care activities was based on the analysis of activities and provisions in the field of health care. During the execution of the project no attempt was made to directly include or link the list of activities to the elements of the classification mentioned above, or to existing international classifications like ISIC/NACE (International Standard Industrial Classification of All Economic Activities/Statistical Classification of Economic Activities in the European Community).

Based on the information available in blocks I to V, the BCP's and the CCP's for the participating countries were constructed, ultimately resulting in the CCP of total health care for the participants.

11.4.1.2 The list of activities in health care and international classifications

The list of activities in health care as used in the CCP project (see Annex 11.1) and in the EUCOMP⁴⁾ project (in a revised form, see Annex 11.2) is not a classification, as it lacks systematic structure, but it could be used "... to provide a common language ... both for the compilation and the presentation of statistics." (NACE, rev. 1, Introduction, page 5). To be effective, the list of activities should be linked into the co-ordinated classifications developed by international organisations, like NACE, and the Statistical Classification of Products by Activity, CPA (see Annex 11.3) or the CPC, Central Product Classification.

NACE is the classification of economic activities "... designed to categorise data that can be related only to the unit of activity." (NACE, rev. 1, Introduction, page 5). The CPA is the product classification of the EU "... designed for categorising products (goods and services) that have common characteristics." (NACE, rev. 1, Introduction, page 5).

There are at least three aspects to harmonising classifications (CPA 1996, Introduction, page 5):

- harmonising classifications of the same type. This means that elements of one classification are comparable with those of another, in a 1:1, 1:n or n:1 relationship (the last two relationships mean a (des)aggregation of one classification compared with the other). The structure of the classification has to be taken into

Stellingen

1. Gezondheid is een relatief begrip dat voornamelijk bepaald wordt door ongezondheid te meten; internationale vergelijkbaarheid is ook een relatief begrip dat voornamelijk bepaald wordt door onvergelijkbaarheid vast te stellen. (Dit proefschrift, hoofdstuk 2)
2. De basis van elke betrouwbare (nationale) statistiek is een systematische registratie van basale data, die een getrouwe weergave van de werkelijkheid vormen. (Dit proefschrift, hoofdstuk 8)
3. Op een lager dan geaggregeerd landelijk niveau (zoals bijvoorbeeld op het niveau van bedrijfstakken) levert het Systeem van Nationale Rekeningen geen uitkomsten die vergelijkbaar zijn tussen landen. (Dit proefschrift, hoofdstuk 3)
4. Vergelijkbaarheid op het terrein van de gezondheidszorg tussen landen kan het beste gerealiseerd worden door specifiek voor dit doel ontworpen classificaties, waarin een strikte scheiding van kenmerken een noodzakelijke voorwaarde is. Het onderbrengen bijvoorbeeld van kenmerken van verstrekkers (zoals de productiemodus) in de classificatie van functies is niet geoorloofd. (Dit proefschrift, zie hoofdstukken 11 en 12)
5. Verschillende typen gebruik vereisen daarop afgestemde datasets. Het gebruik van niet op een bepaald doel afgestemde data kan tot gevolgtrekkingen leiden die niet correct zijn. (Dit proefschrift, zie hoofdstukken 15 en 16)
6. Het aantal klachten en langdurige aandoeningen blijkt een goede benadering te zijn van de door personen ervaren ziektelast. De mate waarin respondenten hetzelfde antwoord geven na een periode van gemiddeld 4 à 5 maanden is minder groot dan de term chronische of langdurige aandoeningen of beperkingen doet vermoeden. (J.L.A. van Sonsbeek, 1996, *Vertel me wat er scheelt: betekenis en methodische aspecten van enquêtevragen naar de gezondheid*)
7. Het algemene gezondheidsoordeel (door de persoon zelf ervaren en gerapporteerd) is nationaal een goed bruikbare en discriminerende maat voor de vergelijkingen van de gezondheidstoestand tussen (bevolkings)groepen. (J.L.A. van Sonsbeek, 1996, *Vertel me wat er scheelt: betekenis en methodische aspecten van enquêtevragen naar de gezondheid*). Bij vergelijking van de uitkomsten tussen landen moet de invloed van landspecifieke kenmerken, die hoofdzakelijk cultureel en historisch bepaald zijn, niet onderschat worden.