KAPITOLA 15

MĚŘENÍ VÝKONU A HODNOCENÍ

Tato kapitola se zabývá měřením výkonu, monitorováním a hodnocením programů. Tyto nástroje a činnosti mají společný cíl: zlepšit řízení programů, zvýšit odpovědnost a zlepšit rozhodovací procesy, protože zpětné informace o výstupech a výsledcích současných politik vlády a vládních programů slouží ke zlepšení tvorby a implementace těchto programů v současnosti i v budoucnosti.

A. Měření výkonu

1. Co je měření výkonu?

Měření výkonu je nástroj, jehož pomocí porovnáváme skutečný vývoj se zadanými záměry a cíli programů, za předpokladu, že jsou strategické cíle známy. Skládá se z následujících činností:

- dokumentování „výrobního procesu“, který je tvořen procesy a činnostmi souvisejícími s přeměnou *vstupů*, což jsou zdroje programu, ve *výstupy*, což je zboží a služby jím vytvořené,

- vyhodnocení *výsledků* – jedná se o široce pojaté ekonomické a sociální změny, které jsou důsledkem politiky nebo programu, a o jejich porovnání s cíli programu.

Měřením výkonu můžeme obecně poukázat na výsledek určité politiky nebo programu, nedostaneme však odpověď na to, proč byl výsledek právě takový, či jaké změny je třeba provést v činnostech nebo ve stanovených cílech programu. Pro tyto účely je nutné provést hloubkové hodnocení programu. Hodnocení programů, jehož přehled uvádí oddíl B, jde za rámec sledování a monitorování ukazatelů výkonu, protože zkoumá způsob, jakým daný program ovlivňuje výsledky. Zatímco měření výkonu se soustřeďuje na efektivnost a účelnost, hodnocení navíc zahrnuje i takové otázky, jako jsou docílené užitky, význam a udržitelnost daného programu.

Měření výkonu obvykle zahrnuje pět následujících hledisek výkonu (OECD, 1994):

- *nákladovou efektivnost,* což je vztah mezi zbožím a službami získanými na základě programu či činnosti (výstupy) a zdroji využívanými k jejich získání (vstupy); měří se náklady na jednotku výstupu. Efektivnost se posuzuje ve srovnání s určitým pevně daným ukazatelem, kterým mohou být např. jednotkové náklady na určitou činnost v předchozím období nebo jednotkové náklady na obdobnou činnost v jiné agentuře či instituci;

- *účelnost*, což je míra, v jaké programy dosáhly očekávaných záměrů nebo výsledků. Účelnost je nejvýznamnějším prvkem hodnoty peněz ve veřejném sektoru. Zboží a služby lze produkovat efektivně z hlediska nákladů vynaložených na jednotku výstupu, ale když nedosáhnou zamýšleného cíle a nepřinesou uspokojení uživatelům veřejných služeb, bude použití zdrojů většinou plýtváním;

- *hospodárnost,* kterou můžeme definovat jako „dosažení odpovídající kvality a kvantity finančních, lidských a hmotných zdrojů v požadovaném čase při co nejnižších nákladech“[1]; lze ji zjišťovat měřením vstupů a jejich porovnáním s normami a standardy;

- *soulad se zákonem.* Agentury musejí postupovat v souladu se zákonem o rozpočtu nebo o rozpočtových přídělech a v souladu s dalšími zákony a nařízeními, např. ve vztahu k řízení hotovostních toků a včasnosti plateb věřitelům. Instituce vybírající daně může mít např. specifické cíle (např. vyjádřené objemem daňových nedoplatků). Měření finančního výkonu však někdy patří spíše do oblasti účinnosti než postupu v souladu se zákonem;

- *kvalitu služeb*. Kvalita služeb v širokém pojetí se týká efektivnosti. Používá se však i v užším smyslu, a to ve vztahu k bezprostředním potřebám uživatelů, jako je včasnost, dostupnost, spolehlivost a plynulé poskytování služeb. V tomto smyslu se týká spíše kvality dodávaných služeb než přínosů služeb samotných. Orientace veřejného sektoru při poskytování služeb na klienty, resp. spotřebitele, je součástí reforem většiny zemí OECD a i mnoha dalších.

„Výkon“ je spojujícím prvkem všech těchto hledisek. Některé z uvedených aspektů jsou na sobě závislé a mohou se dostat do rozporu. Např. je možné zlepšit kvalitu služeb či zlepšit jejich soulad se zákony, ale jen za cenu vyšších nákladů anebo nižší efektivnosti.

2. Měření výkonu a jeho ukazatele

a) Druhy ukazatelů a způsob měření

Výkon lze měřit pomocí různých ukazatelů; přímých, které přímo měří vstupy, výstupy a výsledky, či nepřímých, které se používají jako zástupné měřítko tehdy, když je získání přímého ukazatele obtížné nebo nákladné (např. „pouliční“ cena ilegálních drog jako měřítko pro výsledky protidrogových programů). V praxi, stejně jako v této kapitole, se termíny „ukazatel“ či „měřítko“ používají jako vzájemně zaměnitelné.

Hodnocení výše jmenovaných aspektů výkonu umožňují následující kategorie:

- *Vstupy*. Ukazatele vstupů popisují využití lidských zdrojů, kapitálu, materiálů apod. Vstupy se obvykle vyjadřují jako objem výdajů nebo rozsah pracovní doby. Ukazatele vstupů se týkají *hospodárnosti,* s jakou se zdroje využívají k dosažení výstupů a výsledků.

  Vztáhneme-li ukazatele vstupů jako poměr k výstupům nebo k výsledkům, pak je používáme jako ukazatel *hospodárnosti* či *nákladové efektivnosti*. Do vstupů je třeba zahrnout jak běžné výdaje, tak spotřebu kapitálu. V ideálním případě by se měly vstupy odhadovat podle akruálního principu. Např. analýza programu údržby silnic by měla vzít v úvahu odpisy zařízení, které tvoří významný podíl úplných nákladů. I když běžné náklady nejsou měřítkem úplných nákladů programu, trend těchto nákladů se obecně od trendu úplných nákladů programu příliš neliší (Premchand, 1993).

- *Výstupy*. Výstupy se týkají zboží a služeb zajištěných na základě programu nebo určité činnosti (např. počet kilometrů vybudovaných silnic, počet očkovaných dětí atd.). Ukazatele výstupů se používají ke stanovení *efektivnosti*. Efektivnost lze měřit poměrem vstupů a výstupů, často se vyjadřuje jako počet zaměstnanců nebo délka pracovní doby zaměstnanců na jednotku výstupu a tento ukazatel se nazývá *jednotkové náklady* (např. počet dnů na opravu nebo náklady na opravu jednoho kilometru silnice provedenou uspokojivým způsobem). *Produktivita* se obvykle měří jako poměr objemu výstupů ke vstupům, např. počet přepravených vězňů dělený náklady na jejich přepravu.) Jako zástupný

ukazatel pro měření výstupu se často používají pracovní úvazek nebo ukazatel rozsahu určité činnosti (např. počet provedených inspekcí).

- *Výsledky*. Výsledky odpovídají zamýšleným cílům politik nebo jejich žádoucím záměrům odrážejícím se v dosažených výstupech (zlepšená dostupnost odlehlých území, snížený počet případů určitého onemocnění atd.). Měření výsledků souvisí s *účelností.*

- *Mezivýsledky.* Od nich se očekává, že povedou postupně k zamýšlenému cíli, ale samy o sobě cílem nejsou. Mnoho programů se skládá z posloupnosti mezivýsledků. Např. u programů týkajících se životního prostředí by posloupností mezivýsledků mohlo být: přijatý zákon; počet podniků, které změnily své chování; snížení objemu nebezpečných odpadů a počtu znečišťovatelů. Existují různá označení, která se pokoušejí tyto rozdíly postihnout. Např. „výsledek" lze použít ve stejném smyslu jako „mezivýsledek" či okamžitý výsledek, zatímco „dopad" lze použít současně pro popis okamžitého i konečného výsledku. V nedávno vydané příručce Evropské komise se používá „dopad" k popisu dopadů programu na společnost a „původní dopad" se používá ve smyslu „výsledek" a „dlouhodobý dopad" se pak týká „konečných výsledků".[2]

Kvalitu poskytovaných veřejných služeb měří ukazatele *spotřebitelské spokojenosti* (např. počtem obdržených stížností, z průzkumů, z počtů účastníků). Ukazatele *kvality (poskytnutých) služeb* měří jejich včasnost, dostupnost, spolehlivost a přesnost (např. rychlost reakce opatření, dodržování jízdních řádů, čekací doba v nemocnicích apod.). Kvalita služeb často závisí i na dále vymezených procesech.

K *měření „výrobního" procesu* slouží ukazatele zachycující způsob, kterým se vytváří výstup při dané úrovni zdrojů, dosažená efektivnost a účelnost. Procesy obsahují řetěz činností nebo pracovních postupů, jako např. postupy při veřejných zakázkách, technologické procesy při výrobě zboží a stručné přehledy pro formulaci opatření.

Procesní ukazatele pomáhají při hodnocení výkonu v oblastech, ve kterých se výstupy a výsledky měří obtížně; mají svou hodnotu samy o sobě, zejména při stanovení kvality poskytnutých služeb. V určitých oblastech, v nichž použití ukazatelů výstupu není příliš smysluplné a ukazatele výsledků jsou zase využitelné obtížně, se výkon někdy posuzuje prostřednictvím ukazatelů procesu. Např. stručné přehledy lze použít k hodnocení procesu poskytování poradenských služeb ministrům. V některých oblastech činností veřejného sektoru, jako je např. právo nebo politika, je klíčovým prvkem dobré správy „správný proces".

Jak už jsme naznačili dříve, *dopad* opatření nebo programu se často používá jako synonymum pro výsledek. Někdy však má „dopad" přesnější vymezení. Občas se rozlišuje mezi *hrubým výsledkem* a *čistým dopadem*. Čisté dopady jsou výsledky, které lze plně připsat k určitému programu. Nezahrnují důsledky vnějších efektů programu; odhadují se prostřednictvím metod hodnocení uvedených v oddílu B. Analýza dopadů se týká posouzení důsledků realizace opatření na okolí (např. studie o dopadech na životní prostředí) nebo ukazuje rozsah, v němž program cílenou populaci skutečně ovlivní. Někdy měření dopadu ukazuje, jak výsledek určitého programu ovlivní další programy nebo poslání nějaké organizace.

Sociální ukazatele se používají k posouzení široce definovaného dopadu politik vlády. Obsahují ukazatele vysoce agregované, např. míru dětské úmrtnosti či míru gramotnosti dospělých. Používat sociální ukazatele k posouzení účinnosti určitého programu v praxi není jednoduché. Mohou vznikat problémy například při stanovení příčinné souvislosti (např. snížení dětské úmrtnosti může být důsledkem programu zajištění čisté vody, programu očkování, programu výživy nebo vnějších faktorů, které nejsou výsledkem vládních politik). Ukazatele výsledků programu se navrhují tak, aby se zaměřily na podrobnější aspekty výkonu. Sociální ukazatele přesto poskytují pro rozhodovací procesy užitečné podkladové informace.

Ukazatele výkonu mohou být *kvantitativní* nebo *kvalitativní*. Kvalitativní ukazatele lze transformovat na kvantitativní prostřednictvím průzkumů, respondenčních anket nebo dalších metod, které se používají ke zjištění názorů uživatelů. Např. kvalitu vzdělávání můžeme částečně kvantifikovat na základě zjištění, kolik procent rodičů je „plně spokojeno" se školou, kterou navštěvují jejich děti.

Konečným cílem vládních programů je dosáhnout vytyčených cílů. Samotné stanovení výsledků a nalezení adekvátních ukazatelů může být obtížné. Někdy se výsledky dostaví až po mnoha letech. Program také může být pouze jedním z faktorů, které výsledky ovlivní. Stanovit, v jaké míře lze výsledek připsat danému programu a v jaké míře vlivům jiným, je náročné. Pro posouzení úspěchu programů jsou ukazatele výsledku významnější než ukazatele výstupu, obvykle se však snadněji definují ukazatele výstupu a rovněž jejich měření je snazší.

V některých oblastech lze použít ukazatele výstupu jako náhradu za ukazatele výsledku. Ke stanovení efektivnosti programu výstavby silnic lze např. použít počet nově vybudovaných kilometrů silnic, což je výstup tohoto programu. V dalších oblastech, zejména v oblasti sociální, mohou být nicméně výsledky programu velmi vzdálené od jeho výstupu, takže ukazatele výstupu ukazatelem výsledku vždy spolehlivě nahradit nelze. Např. vyšší počet návštěv u lékaře nutně neznamená snížení nemocnosti. Pro ty, kdo rozhodují, může být užitečné zajištění zpětné vazby formou měření mezivýsledků.

b) Výkon jako relativní pojem

Výkon je jenom relativní pojem. Z podstaty věci vyplývá, že hodnocení efektivnosti vyžaduje, abychom porovnali ukazatele výsledků nebo výstupu s cílem programu. Pro posouzení toho, zda výsledky jsou dobré, špatné, nebo neurčitelné, bychom měli prakticky každý ukazatel výkonu srovnávat s nějakou základnou nebo standardem. Výkon se pak měří ve srovnání:

- s tím, čeho bylo dosaženo v minulosti. Časové řady údajů jsou velmi užitečné, neberou však ohled na změny v efektivnosti a v produktivitě, které jsou důsledkem technologických faktorů;

- s tím, čeho dosahují srovnatelné programy nebo organizace, nebo s existujícími národními či mezinárodními standardy v dané oblasti. Činnosti jiných organizací poskytují užitečné referenční ukazatele, problém však spočívá v nalezení organizace skutečně srovnatelné;[3]

- s cíli stanovenými v rozpočtu nebo s jinými opatřeními vlády.

Srovnávat bychom však měli pouze srovnatelné. Je proto nutné, abychom správně definovali ukazatele a základ pro srovnání. Např. při porovnávání výsledků zkoušek na různých školách může být vhodné očistit hrubé ukazatele (např. počet zkoušek, které udělal jeden student) o různé vlivy, např. o rozdíly v sociálním původu studentů.

Ukazatele výkonu sice mohou hrát užitečnou roli, ale je třeba je používat s určitou dávkou opatrnosti. Je nutné systematicky zkoumat jejich význam a interpretaci, protože – pokud se tak neděje – mohou vážně narušit fungování organizace, chování manažerů a zaměstnavatelů (viz Likierman, 1993b). Zaměříme-li se např. striktně pouze na malý okruh ukazatelů výkonu bez diskuze o jejich interpretaci, můžeme sice úspěšně dosáhnout některých cílů, ale tento postup může odvést naši pozornost od širších organizačních cílů a záměrů.

3. Funkce měření výkonu

a) Odlišné systémy měření pro odlišné účely

Rozvoj a implementace měření výkonu by měly být uzpůsobeny místním podmínkám a záměrům. Podstata měření výkonu se liší v závislosti na odpovědnosti těch, jejichž výkon se měří, a na požadavcích těch, kteří tyto informace používají. Na operační úrovni by se měly ukazatele týkat takových aspektů, jako je řízení zdrojů a výrobních procesů. Na vyšší úrovni řízení by se informace měly týkat efektivnosti programů, aby byl k dispozici dostatek informací pro rozhodování o opatřeních a alokaci zdrojů.

b) Učení se organizaci a řízení programů

Měření výkonu je užitečné pro hodnocení účinnosti administrativních a organizačních postupů. Lze je použít pro zlepšení provozní efektivnosti celé organizace (např. ministerstva nebo agentury); použít je mohou i jednotliví vedoucí pracovníci v rámci organizace k hodnocení a zlepšení výkonu odboru, oddělení nebo dílčího útvaru dané organizace.

Měření výkonu může být užitečným nástrojem pro řízení nárokových programů a investičních programů a projektů. Informace o výkonu může přispět jak ke zjištění toho, zda jsou programy implementovány v souladu s jejich cíli, tak pro přípravu programů nových. Např. v oblasti výstavby silnic může být měření výkonu, které zahrnuje takové aspekty, jako je mobilita/dostupnost, dopravní tahy, bezpečnost a dopady na životní prostředí, užitečné pro přípravu a pro dohled nad realizací programů.

c) Výkonnostní smlouvy a dohody

Měření výkonu lze použít jako nástroj pro posílení odpovědnosti vedoucích pracovníků. Řízení orientované na výsledky usiluje o propojení výkonnosti vedoucích pracovníků s explicitními nebo implicitními smlouvami, které obvykle výkonnostní cíle zahrnují. Teoreticky by smlouvy měly obsahovat jak postihy, tak odměny. Ve většině případů však smlouvy obsahují pouze ustanovení o odměnách, jako např. možnost vedoucích pracovníků získat část docílené úspory nebo úsporu celou nebo možnost pružně přizpůsobovat rozsah vstupů (např. počet zaměstnanců) nebo vázání výše odměn a platů vyšších vedoucích pracovníků a vedoucích jednotlivých útvarů na zvolený výkonově definovaný prvek.

Výkonnostní hlediska v těchto smlouvách jsou nezbytně užší, než když se hodnotí celý program. Vedoucí pracovníky je možné např. činit odpovědnými za program očkování a podle toho je odměňovat. Je však obtížné, aby nesli odpovědnost i za výsledek ve formě lepšího zdravotního stavu, protože ten bude záviset i na faktorech mimo jejich dosah, např. na kvalitě vody nebo změně ve spotřebě tabákových výrobků. Měření výkonu, které se používá pro vymezení odpovědnosti a kontroly, se nezbytně zaměřuje spíše na vstupy a výstupy než na výsledky programů.

I v tom případě, že jsou rozpočtové prostředky poskytovány na základě výkonnostní smlouvy nebo dohody, může být spojování výkonu a alokace zdrojů jen nepřímé. V mnoha případech uvolňování rozpočtových prostředků nebo poskytování dotací není podmíněné dosažením požadovaných výsledků výkonu.

Smlouvy nabízejí způsob, jak doplnit veskrze formální podstatu cílů a ukazatelů, protože kladou větší důraz na vztahy v rámci řízení než na pouhé sbírání kvantitativních informací o úspěších (Trosa, 1996a). Zevšeobecnění smluvního přístupu a jeho uplatnění na více než několik málo programů nebo organizací by však bylo v současné době ve většině reformních jen velmi obtížné.

Platový systém a odměňování orientované na výkon spojují některé prvky odměňování se specifickým rozsahem činnosti nebo výstupu a mohou přispět ke zlepšení výkonu. V reformních zemích by tyto systémy mohly usnadnit přechod od kultury příkazů ke kultuře orientace na zákazníka. Při úvahách o zavedení tohoto systému je však nutná značná obezřetnost. Výběr relevantních ukazatelů je ošidnou záležitostí. Činit zaměstnance odpovědné za věci, které nemají plně pod kontrolou, je sporné. Zaměřit se na výsledky, které lze přímo připsat úsilí zaměstnanců a vedoucích pracovníků, u nich může krátkodobě vyvolat pozitivní reakci podnícenou odměnou, ale může se tak dít na úkor činností, které vedou k dosažení širších cílů programů nebo organizace. To, zda systém odměňování orientovaný na výkonnost skutečně výkon zlepší, je diskutabilní. Teorie motivací zdůrazňují spíše vnitřní stimuly (např. v některých zemích je to pracovní místo samo o osobě nebo etika služeb poskytovaných veřejných sektorem) než podněty vnější (peníze a výhody). Řízení lidských zdrojů orientované na výsledky může navíc vést k nežádoucím výsledkům v těch případech, kdy dominuje politické vměšování a protekcionizmus; za takových podmínek budou politikové a vedoucí pracovníci směřovat spíše k odměňování „svých“ lidí než lidí s nejlepším výkonem.

d) Odpovědnost vůči okolí

V některých členských státech EU, které sledují zvýšení odpovědnosti veřejného sektoru vůči parlamentu a daňovým poplatníkům, se informace o výkonu publikují; tím se také usnadňuje kontrola využívání peněz. Ve Velké Británii se např. požaduje, aby agentury v ročních výkonnostních smlouvách uzavíraných s příslušným ministrem a ve svých ročních zprávách poskytovaly údaje o výkonu za několik let tak, aby je bylo možné v čase srovnávat. V mnoha zemích dává nejvyšší kontrolní orgán připomínky ke způsobu měření výkonnosti a k jeho vhodnosti, ve většině případů však samotné výsledky nekomentuje.

Při úvahách o zpřístupnění informací o výkonu široké veřejnosti je zapotřebí určité obezřetnosti. Na jedné straně může být zveřejňování užitečné z hlediska kontroly a z hlediska výchovného, protože přispívá k etablování určité kultury výkonu. Na druhé straně zde však existuje riziko nesprávného použití informací v případě, že se nerespektují jeho omezení; existuje riziko demotivace zaměstnanců vystavených nespravedlivé kritice. Zveřejnění informací o výkonu podporuje konkurenci mezi obdobnými organizacemi, a tudíž i efektivnost. Může však také přispět k prohloubení odstupu mezi těmi, kteří podávají dobrý výkon, a těmi, kteří podávají výkon špatný. Týká se to zejména oblasti vzdělávání a zdravotní péče, kde se takto může vyvolat přesun zdrojů nebo sociálně lépe postavených studentů nebo pacientů ke školám a nemocnicím, o nichž se říká, že fungují nejlépe.

4. Měření efektivního výkonu

a) Potřeba obezřetnosti

Používat měření výkonu jako nástroj řízení výkonu (např. pro formulaci smlouvy) nebo jako prostředek pro vyjádření veřejné a politické odpovědnosti může být nebezpečné. Zkušenosti centrálně řízených zemí ukazují, že zavedení norem a standardů vytváří situaci, kdy úředníci velmi trvají na dosažení specifických cílů. Pak mohou opomíjet kvalitu a nezbytně nutné propojení cílů a samotnému výkonu nevěnují zaslouženou pozornost. Pokud se standardy dosahují bez ohledu na dostupné zdroje, dojde pravděpodobně k oslabení fiskální kontroly. Standardy a ukazatele bychom neměli považovat za nezměnitelné.

Tyto problémy nejsou spojeny výlučně s centrálně plánovanými ekonomikami a příklady nežádoucích důsledků využívání (nebo zneužívání) ukazatelů výkonu v tržních ekonomikách jsou četné. „Zákon nezamýšlených důsledků“ říká, že pokusy změnit chování mohou vést k nezamýšlenému chování, které může být v rozporu s cíli a záměry politik nebo programů. Stanovuje-li se např. dotace nemocnicím podle délky seznamu čekajících pacientů, stimuluje to vedoucí pracovníky a lékaře k tomu, aby nechávali co nej-

déle čekat případy, které nejsou kritické, a zabývali se jinými záležitostmi (tj. vyšší kvalitou péče o některé a špatnou pro ostatní); v případě, že se posuzuje výkon pouze na základě počtu ošetřených pacientů, může dojít ke zhoršení celkové kvality lékařské péče.

V závislosti na způsobu, jakým se ukazatele výkonu stanoví a použijí, mohou vznikat následující rizika:[4]

- *tunelové vidění* neboli důraz pouze na aspekty výkonu, které lze kvantifikovat, a nikoli na ty, které kvantifikovat nelze;
- *fixace na ukazatele*, resp. soustředění se na to, co lze měřit, a nikoli na služby, které se mají poskytovat;
- *krátkodobost*, věnování nedostatečné pozornosti oprávněným dlouhodobým cílům; volí se tak *suboptimální řešení* čili dosahování nižší kvality služeb, protože dochází k většímu soustředění se na úzce definované činnosti a méně pozornosti se věnuje široce vymezeným cílům organizace;
- *špatná reprezentativnost nebo záměrné zkreslování údajů* a špatná interpretace nebo nekritické přijímání výsledků měření výkonu;
- *strategické řízení chování* včetně úmyslně dosahovaného nižšího výkonu, aby bylo možné jednoduchým způsobem plnit dané požadavky;
- *nepružné usilování o dosažení definovaných cílů*;
- *demoralizace*, resp. ztráta důvěry a angažovanosti pracovníků, kteří poskytují služby, se považuje za méně významný aspekt než samotný obsah měření výkonu.

b) Kritéria systémů měření dobrého výkonu

Pro stanovení ukazatelů výkonu lze použít některá všeobecná vodítka a kritéria, která slouží k tomu, abychom se výše uvedeným úskalím vyhnuli (Shand, 1998):

- *relevance a užitečnost*. Ukazatele by se měly řádně definovat ve vztahu k danému programu a měly by odrážet jeho hlavní cíle a záměry. Výkon vedoucích pracovníků by se měl měřit pouze v těch oblastech, které mohou ovlivňovat;
- *jednoznačnost a srozumitelnost*. Ukazatele výkonu by měly být jednoduché, dobře definované a pro uživatele snadno srozumitelné;
- *nákladová efektivnost*. Ukazatele výkonu by se měly zavádět s přijatelnými náklady. Náklady na sběr údajů k zavedení ukazatelů výkonu a na řízení tohoto systému je třeba odhadovat realisticky a poměřovat je s očekávanými výsledky;
- *schopnost monitorovat výsledky*. Jak jsme již poznamenali dříve, výkon je pouze relativní pojem a měření je nutné aplikovat konzistentně v čase i mezi organizačními jednotkami tak, abychom ho mohli systematicky porovnávat.

5. Referenční ukazatele

Referenční ukazatele, které se používají jak v soukromém, tak ve veřejném sektoru, slouží pro srovnání výkonu jedné organizace se standardem, který může být vyjádřen absolutně nebo relativně k výkonu dané organizace. Lze je použít k:

- posouzení výkonu ve srovnání s definovaným standardem nebo standardy,
- odhalení oblastí, které je třeba zlepšit,
- identifikaci činností, které se provádějí lépe v jiných organizacích,
- testování toho, zda opatření přijatá ke zlepšení efektivnosti a účelnosti programů jsou úspěšná.

Existují dva hlavní způsoby použití referenčních ukazatelů ve veřejném sektoru:

- *referenční ukazatele procesů* se aplikují na procesy a činnosti související s přeměnou vstupů na výstupy. Jedná se o stanovení referenčních ukazatelů pro danou organizaci, a to buď ve vztahu k procesům uplatňovaným v jiných organizacích, nebo ve vztahu k procesům definovaným jako standard;
- *referenční ukazatele výsledků* se aplikují na skutečné výsledky (výstupy). Jedná se o srovnání skutečného výkonu různých organizací, které využívají ukazatele výkonu, nebo srovnání skutečného výkonu s určitými standardy výkonu.

Tyto dva hlavní druhy referenčního srovnávání (procesového a výsledkového) se v stále větší míře používají jako vzájemně se doplňující metody, které se zároveň vzájemně posilují. Použití výsledkových referenčních ukazatelů lze např. použít k identifikaci odchylek ve výsledcích, zatímco pomocí procesových referenčních ukazatelů lze vysvětlit, proč tyto odchylky existují. Použití procesových referenční ukazatelů bez použití ukazatelů výsledkových má tendenci stát se hodnocením směřujícím dovnitř organizace; to vede k soustředění se na zlepšování procesů pro ně samé bez sledování toho, zda se něco změní pro uživatele veřejných služeb a pro další zúčastněné subjekty.

Referenční ukazatele lze používat jako nástroj hodnocení a plynulého rozvoje. Jsou vhodné pro značný počet manažerských technik jako např. pro řízení celkové kvality a procesového reingeneeringu a užívají se také pro srovnávání výkonu a pro hodnocení programů. Reformní země by měly zvážit možnost využívání referenčních ukazatelů k hodnocení vybraných případů. Vzhledem k tomu, že využívání referenčních bodů může být náročné na čas a zdroje, je vhodné se zaměřit nejprve jen na několik klíčových organizací nebo procesů.

Box 15.1 ukazuje některé příklady výkonově orientovaných ukazatelů používaných v sektoru zdravotní péče ve Velké Británii.

Box 15.1 VÝKONOVÉ UKAZATELE VE ZDRAVOTNICTVÍ VE VELKÉ BRITÁNII

Oblasti a kategorie	Ukazatele
I. Zlepšení zdravotního stavu	
Celkový zdravotní stav obyvatelstva, který odráží sociální aspekty a aspekty životního prostředí a individuální chování a dále péči poskytovanou NHS[1*] a dalšími agenturami. Všechny příčiny úmrtí (lidé ve věku 15–64 let).	1. Počet zemřelých bez ohledu na příčinu úmrtí (věk 15–64) 2. Počet zemřelých bez ohledu na příčinu úmrtí (věk 65–74) 3. Evidence výskytu rakoviny
II. Spravedlivý přístup	
Dostupnost lékařského zákroku	1. Počet operací
Přístup ke službám plánovaného rodičovství	2. Počet početí u dívek ve věku 13–15 let
Přístup k zubnímu lékaři	3. Lidé registrovaní u zubního lékaře-člena NHS
Přístup k prevenci zdraví	4. Včasná diagnostika rakoviny
Přístup k sociálním službám	5. Návštěvy obvodní sestry
III. Efektivní poskytování vhodných služeb zdravotní péče	
Zlepšení zdraví/prevence nemocí	1. Prevence nemocí a zlepšení zdraví 2. Včasná diagnostika rakoviny
Vhodnost operativních zákroků	3. Nevhodně provedené operační zákroky 4. Počet operací
Řízení primární péče	5. Péče o akutně nemocné 6. Péče o chronicky nemocné 7. Duševní zdraví v primární péči 8. Hospodárné předepisování léků
Dodržování standardů	9. Počet propuštěných z nemocnice
IV. Efektivnost	
Maximální využití zdrojů	1. Počet případů za den 2. Délka pobytu v nemocnici 3. Jednotkové náklady 4. Předpisy generických léků

(pokračování)

Box 15.1 VÝKONOVÉ UKAZATELE VE ZDRAVOTNICTVÍ VE VELKÉ BRITÁNII

(pokračování)

Oblasti a kategorie	Ukazatele
V. Zkušenosti pacientů ošetřených v NHS	
Dostupnost	1. Pacienti, kteří čekají déle než 2 hodiny při akutním onemocnění
	2. Pacienti, u nichž byla zrušena operace z důvodů jiných než medicínských v den přijetí nebo jeden den po přijetí
Koordinace a komunikace	3. Odložené propuštění z nemocnice u lidí nad 75 let
	4. První další termín pro ambulantní pacienty, kteří se nedostavili k ošetření
Čekací doba	5. Návštěva ambulantních pacientů do 13 týdnů po písemném oznámení o přijetí do nemocnice
	6. Pacienti přijatí do 3 měsíců po rozhodnutí o přijetí
VI. Zdravotní stav jako výsledek péče NHS	
Úspěšnost při snižování rizika	1. Počet početí u dívek ve věku 13–15 let
Úspěšnost při snižování nemocnosti, poškození zdraví a komplikací při léčbě	2. Zkažené, chybějící nebo plombované zuby u dětí ve věku 5 let
	3. Nemoci, kterým se dalo vyhnout
	4. Nepříznivé okolnosti/komplikace při léčbě
Úspěšnost při optimalizaci funkcí a zlepšení kvality života pacientů a jejich ošetřování	5. Naléhavý příjem do nemocnice u lidí nad 75 let
	6. Počet opětovného přijetí pro psychiatrické onemocnění
Úspěšnost při snižování předčasných úmrtí	7. Kojenecká úmrtnost
	8. Míra přežití při rakovině prsu a děložního hrdla
	9. Odvratitelná úmrtí
	10. Předčasná úmrtí v nemocnici
VII. Příklad rakoviny prsu	1. Podchycení rakoviny
	2. Podchycení rakoviny plus míra jejího rozvinutí při první diagnostice
Úspěšnost NHS při snižování nemocnosti, při poškození zdraví a při komplikacích při léčbě	3. Výskyt komplikací, kterým lze předejít – opětovný výskyt, komplikace při léčbě atd.
Úspěšnost NHS při obnově funkcí a při zlepšování kvality života pacientů	4. K měření se používá dotazník k hodnocení pacienty nebo jiné vhodné prostředky
Úspěšnost NHS při snižování počtu předčasných úmrtí	5. Pětileté přežití
	6. Pětileté přežití upravené vzhledem k věku a stadiu nemoci

Poznámka: NHS/v NHS = (v zařízeních) britské Národní zdravotní služby

B. Hodnocení programů

1. Definice a cíle[5]

a) V čem spočívá hodnocení programů?

Hodnocení programů se zaměřuje na porovnání výsledků programů s jeho cíli. Často se také používá označení „hodnocení politik"; to však má širší význam, protože se může týkat i několika programů; lze také použít označení „regulatorní rámec", „analýza vztahů mezi programy a nařízeními" ad. Mnohé z problémů, které dále popisujeme, se týkají také hodnocení politiky. Některé země však hodnocení programů od hodnocení politiky neodlišují.[6]

Hodnocení programů může probíhat v různých etapách jejich realizace:

- Hodnocení *ex post* se provádí poté, co program již nějakou dobu funguje. Slouží k analýze jeho efektivnosti a k posouzení jeho celkového významu. Tato hodnocení se obvykle používají jako podpůrný nástroj rozdělování rozpočtových prostředků nebo jako nástroj zvýšení odpovědnosti za jejich užití. Očekává se, že odpovědí i na otázky týkající se výsledku a celkového významu programu.

- *Průběžné* hodnocení se obvykle provádí v průběhu realizace programu. Jeho smyslem je podpora a zlepšení řízení a implementace programu. Důraz se klade na otázky související s fungováním programu.

b) Cíle

Cílem hodnocení je zlepšit rozhodování a alokaci zdrojů, a to tím, že se zjišťují spolehlivé údaje o důsledcích politik a programů. Hodnocení programů lze využít k následujícím účelům:

- jako *pomoc při alokaci zdrojů a při identifikaci žádoucích změn politik.* Hodnocení poskytuje informace o důsledcích existujících politik. Pomáhá těm, kdo o politikách rozhodují, při stanovení významu veřejných programů a při identifikaci oblastí, v nichž může být nezbytná změna politiky nebo změna v alokaci zdrojů mezi různými programy;

- *ke zlepšení řízení programu a institucionálních postupů.* Jak jsme poznamenali dříve, mechanizmus zpětné vazby přispívá k procesu učení se těch, kteří program řídí nebo jej realizují, a lze ho použít také ke zlepšení institucionálního zajištění;

- *k prohloubení odpovědnosti vůči veřejnosti.* Hodnocení může zlepšit transparentnost a odpovědnost tím, že poukáže na dopady vládních politik.

c) Hodnocení, monitorování a kontrola

Hodnocení se liší od ostatních mechanizmů zpětné vazby, jako je monitorování a měření výkonu, protože je obecně prováděno jako samostatná akce, během níž se shromažďují informace, které zasahují do větší hloubky. Zatímco měření výkonu se soustřeďuje na nákladovou efektivnost a účelnost a často i na další užší otázky, hodnotící studie posuzují také to, zda program reaguje na potřeby a sociální a ekonomické problémy, které má řešit. Hodnotící studie často obsahují podrobný přehled přínosů a příčinných souvislostí výsledků, zatímco měření výkonu se zabývá pouze hrubě vymezenými ukazateli výstupů a výsledků.

K úspěchu hodnocení může užitečnými informacemi přispět pravidelné monitorování a měření výkonu. V některých programech jsou např. již předem zabudovány etapy, v nichž se sleduje úspěšnost plnění určitých cílů a záměrů; to pak slouží jako „kotva" podrobnějšího zkoumání úspěchů a selhání programů.

Hodnocení a vnější kontrola jsou historicky oddělené funkce, které vykonávají samostatné instituce. Kontrola je úzce spojena s dohlížecí funkcí parlamentu; u organizací veřejného sektoru pak zvyšuje význam dodržování právního rámce a jejich odpovědnost za výsledky; nezávislost nejvyšší kontrolní instituce je, jak jsem probrali v kapitole 14, stěžejní. Hodnocení je pak obecně činnost, za kterou je odpovědná exekutiva. I když se hodnocení stále více využívá ke zvyšování odpovědnosti, je především nástrojem posilování programového managementu a podpory rozhodovacích procesů. V praxi se však hranice mezi hodnocením a kontrolou stírají, protože tradiční finanční kontrola se doplňuje o kontrolu odpovědného využívání peněz daňových poplatníků – což jsou metodologicky podobné problémy.

2. Klíčové otázky hodnocení

a) Logika programů

Klíčovou otázkou hodnocení programů je zkoumání jejich „vnitřní logiky" (tj. jejich základní filozofie s cílem zjistit, v jakém rozsahu se dosahuje cílů a záměrů programů); některé manuály o hodnocení odkazují na „teorii programů".

Programy se vždy připravují se zřetelem k určitému okruhu potřeb. Těmito potřebami jsou sociálně-ekonomické cíle a otázky, které se program pokouší řešit, vyjádřené z hlediska specifické cílové populace. Podle logického rámce, který přijala Evropská komise[7], lze tyto cíle rozdělit na:

- *obecné cíle*, které se vyjadřují ve formě konečných výsledků,
- *specifické cíle*, které se vyjadřují ve formě mezivýsledků,
- *operační cíle,* které se týkají plánovaných vstupů a výstupů.

Strukturovaný přístup by měl obsahovat následující kroky: (1) popis programu, (2) vyjasnění cílů a potřeb, o jejichž naplnění program usiluje, (3) identifikaci možných kauzálních vztahů mezi působením programu a jeho důsledky, (4) identifikaci možných výsledků, které lze hodnotit (mezivýsledky anebo konečné výsledky), (5) identifikaci ukazatelů výsledků a kritérií, která se použijí ke stanovení efektivnosti, a (6) identifikaci faktorů, které mohou výsledky ovlivnit. Box 15.2 ukazuje příklad takového logického rámce pro hodnotící studii, který vychází z metodologie Evropské komise.

Obdobně jako u systému měření výkonu popsaného v oddílu A musíme i zde určit relevantní ukazatele. Mohou přitom vzniknout problémy obdobné těm, které jsme probrali v předchozím oddílu. Zvlášť významný pro hodnocení výsledků programu je problém týkající se stanovení toho, zda a v jakém rozsahu daný program vedl k pozorovaným důsledkům. Hodnocení programu vyžaduje porovnat výsledky se záměry uvedenými v programu nebo se specifickými referenčními ukazateli. Jsou-li cíle a záměry programu při formulaci rozpočtové politiky jasně určeny, je jejich hodnocení výrazně jednodušší.

Box 15.2 PŘÍKLAD PŘÍSTUPU K HODNOCENÍ POMOCÍ LOGICKÉHO RÁMCE

	Logika zásahu	Ověřitelné ukazatele
Konečné výsledky	Moderní, soukromý a transparentní bankovní sektor, pro který je charakteristická zdravá proporce mezi rizikem a odměnou	Rating výkonu komerčních bank
Mezivýsledky	Restrukturalizace a reorganizace komerčních bank v životaschopné instituce a jejich příprava na následnou privatizaci	Počet privatizovaných bank
		Míry obezřetnosti a výkonové ukazatele pro banky, podíl špatných úvěrů, úroveň rezerv a oprávek atd.
		Přijaté záruky
		Implementace plánu nápravných opatření
		Organizační struktura bank
Operační cíle	Vzdělaní vedoucí pracovníci a zaměstnanci	Fungující vybavení
Výstupy	Manuály	Existence restrukturalizačních plánů a jejich prodiskutování
	Nová organizační struktura	Organizační struktura akceptovaná vedením
	Instalovaná zařízení	
	Plány korektivních kroků	Počet proškolených zaměstnanců
	Restrukturalizační akční plány	
Vstupy	Technická pomoc (školení, předprivatizační pomoc atd.)	Doba trvání technické pomoci konzultantů a dalších smlouvou vázaných pracovníků

Zdroj: Komise Evropských společenství, hodnocení bankovních programů Phare, 1998

b) Další klíčové otázky

Schéma 15.1 ukazuje logický rámec a hlavní otázky, na které musí studie hodnotící výkon poskytnout odpověď. Tyto klíčové otázky můžeme zařadit do následujících kategorií:

- *Stálost významu.* Rozsah, ve kterém je program relevantní z hlediska priorit vlády. V jakém rozsahu jsou cíle a mandát programu stále relevantní? Jsou činnosti a výstupy konzistentní s mandátem programu a jsou přijatelným způsobem navázány na jeho cíle a další zamýšlené výsledky?

- *Užitečnost.* Jak vyhovuje uvažovaný dopad programu potřebám cílové populace?

- *Udržitelnost.* V jakém rozsahu mohou všechny pozitivní změny, jež jsou důsledkem programu, přetrvat i po jeho skončení?

- *Efektivnost.* Jak úsporně a efektivně byly různé vstupy přeměněny na výstupy?

- *Účelnost.* Byly záměry programů dosaženy? Jaké nastaly zamýšlené a nezamýšlené přínosy pro klienty a jaké širší důsledky vznikly realizací programu? Doplňuje tento program jiné programy, kopíruje je nebo se s nimi v některých oblastech překrývá?

- *Nákladová efektivnost.* Existují alternativní řešení nebo programy, které by dosáhly žádoucí cíle a zamýšlené výsledky s větší nákladovou efektivností? Existují efektivnější způsoby provedení probíhajícího programu?

Kromě těchto otázek by měla hodnotící studie také posoudit, zda výsledkem programu nejsou negativní efekty nebo neefektivnosti, které se v mikroekonomii nazývají „efekt umrtvené váhy“, „efekt přesunu“ či „efekt substituce“. „Efekt umrtvené váhy“ se vztahuje k těm přínosům programu, které by se byly realizovaly, i kdyby se program neuskutečnil. Např. rekvalifikačním programem pro dlouhodobě nezaměstnané mohou projít i lidé, kteří by se rekvalifikovali i bez existence tohoto programu. „Efekt přesunu“ a „efekt substituce“ se používají k popisu situace, ve které je přínos programu pro určitého jedince, skupinu nebo oblast dosažitelný pouze za cenu ztráty pro jiné jedince, skupiny nebo oblasti. U programu dotací na podporu cílené zaměstnanosti např. může dojít k substituci tehdy, když v podniku, který má z tohoto programu prospěch, najdou práci jedinci, kteří jsou do tohoto programu zahrnuti, na úkor jedinců, kteří do něj nespadají. K přesunu dojde tehdy, když podnik, který má prospěch z podpor zaměstnanosti, získá zakázku na úkor ostatních firem, které se tohoto programu nezúčastní. Tvorba pracovních míst v určité firmě může být tudíž kompenzována ztrátou pracovních míst ve firmách jiných.

3. Příprava hodnotící studie

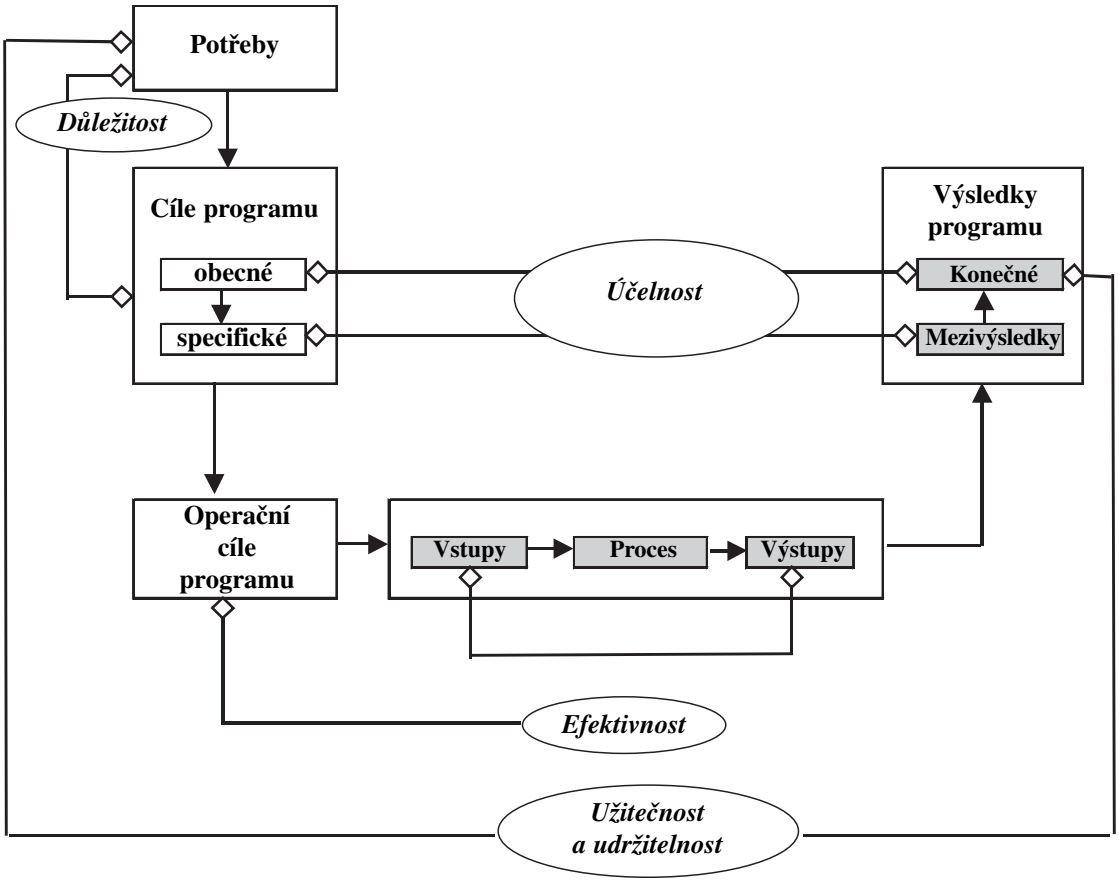
Existují dvě hlavní fáze zpracování hodnotící studie:

- *projekt* – identifikace hlavních otázek a problémů, kterými je nutné se zabývat, a vývoj metodologie pro sběr a zpracování informací,

- *implementace* – sběr a analýza údajů, vypracování první verze zprávy, v níž jsou uvedena zjištění studie a její doporučení.

Většina návodů, jak zpracovat hodnotící studie, zdůrazňuje význam fáze přípravy. Klíčové kroky při přípravě hodnotící studie jsou:

Schéma 15.1 **LOGIKA PROGRAMU, MĚŘENÍ VÝKONU A JEHO HODNOCENÍ**



- *stanovení cílů hodnocení.* Důležitou prvotní otázkou je: Proč se má hodnocení provést? Pro zlepšení řízení? Pro lepší rozhodování o politikách? Pro zvýšení odpovědnosti?

- *vymezení rozsahu hodnocení.* Zahrneme-li do programu několik cílů a cílových skupin, může být někdy nákladově efektivní omezit hodnocení pouze na určité aspekty programu;

- *stanovení otázek, které je dobré si položit.* Pokud je smyslem studie hlavně zlepšit řízení, měla by se zaměřit na prověření implementace programu a na prověření zajištěnosti služeb. Pokud hodnotíme otázku odpovědnosti, měla by se studie zaměřit na efektivnost programu. Aby studie poskytla zpětnou vazbu těm, kteří rozhodují, měla by zahrnovat analýzu nákladů a přínosů a také stanovit, zda byl program užitečný a zda má smysl v něm pokračovat;

- *zformování logiky programu;*

- *stanovení referenčních ukazatelů.* Hodnocení znamená určit „hodnotu" programu. To zahrnuje posouzení míry, v níž je výkon programu „dobrý" nebo „špatný". Jen předem stanovené a transparentní referenční ukazatele mohou zajistit, že se posouzení „hodnoty" programu nestane arbitrární záležitostí;

- *příprava analytického postupu.* Zahrnuje definování postupu hodnocení, metodiky sběru údajů a postupů jejich analýz. Dále by se měl připravit popis metody, jež se použije ke sběru informací a k vypracování závěrů o výsledcích, které jsou důsledkem programu. Rámec postupu závisí na druhu potřebných informací a na druhu analýzy, která se jimi bude zabývat;

- *zajištění dostatečného objemu dostupných informací.* Pro většinu práce na hodnotící studii bude prvním zdrojem informací monitorovací systém. Tyto informace je však nutné doplnit zkoumáním dalších statistických zdrojů, dotazníků a uživatelských průzkumů apod.;

- *příprava plánu práce a odhad nákladů na hodnocení;*

- *určení kritérií pro hodnocení;*

- *výběr hodnotitele.*

4. Záměr hodnocení

a) Existuje zlaté pravidlo?

Existuje několik způsobů, jak hodnotící studii zpracovat. Dostatečný přehled o tom, jak lze hodnocení provádět, je nezbytný pro vymezení vhodné strategie hodnocení. V literatuře o hodnocení najdeme hodně diskuzí o výhodách a nevýhodách určitých metod a přístupů k hodnocení. Jak však Cheminski (1995) říká:

„Výběr metod, nástrojů a údajů pro hodnocení závisí více na otázce, proč se hodnocení provádí, než na kvalitě určité metody. Vyžaduje odpověď na tuto otázku, zda má jít spíše o popis, nebo zda se vyžaduje zkoumání kauzální souvislosti. V popisné studii se např. nemusíme zabývat problémem experimentálních metod, protože ty jsou k tomu nevhodné. Pro studii, která zkoumá vztah příčiny a následku, je možné použít řadu různých metod v závislosti na otázce, kterou řešíme. Pokud se soustředíme více na „problém“ než na metodu, nabízí se – vzhledem ke slabým místům jednotlivých metod – komplementarita metod a opakování postupu. Povrchnost ankety zmírňujeme např. tím, že ji doplníme o případové studie, analýzu časových řad „zlidštíme“ tím, že provedeme anketu nebo sérii rozhovorů, a tím, že integrujeme hodnocení procesu do studie o výsledku. To, že se zaměřujeme spíše na otázku než na metodu, nám poněkud rozvazuje ruce při výběru metody a vede opět ke zdůraznění významu plurality problému a metody, což není při hodnocení nedávné minulosti vždy snadné...“

Nebo jak lapidárně říká průvodce hodnocením EK (Komise Evropských společenství, 1997): „Zlaté pravidlo: neexistuje žádné zlaté pravidlo.“

Dále uvádíme stručný přehled možných postupů hodnotících studií a metod sběru údajů a analýzy. Na toto téma existuje nepřeberné množství literatury, ve které lze nalézt přehledy hodnocení, které jsou techničtější a podrobnější.

b) Experimentální nebo kvaziexperimentální postup hodnocení[8]

Hodnotit program vyžaduje porozumět tomu, co program přináší. Z tohoto hlediska je běžným postupem posuzovat přínosy programu v porovnání s tím, co by bylo, kdyby se program neuskutečnil, tj. zjišťovat hypotetické skutečnosti „co by bylo, kdyby bylo“. Protože tento hypotetický výsledek nemůžeme zkoumat přímo, musí hodnotitel aplikovat postupy pro identifikaci hypotetických dopadů a porovnat je s dopady programu, které měříme.

Experimentálně pojatá hodnotící studie má výhodu v tom, že umožňuje získat výsledky s vysokou mírou spolehlivosti. Používá postupy a techniky, které známe z přírodních věd, ke zkoumání veřejných programů. Experimentální postup spočívá např. v tom, že vybereme náhodně dvě skupiny těch lidí, na které je program cílen; obě skupiny by měly ve všech ohledech vykazovat shodné charakteristiky. Jedna skupina je tzv. experimentální či „ošetřená programem" a druhá, tzv. kontrolní, stojí mimo program. V řádně vedeném experimentu lze rozdíly ve výsledcích získaných od těchto dvou skupin připsat důsledkům programu nebo politiky.

Výsledky získané pomocí studií postavených na náhodném výběru mají tu přednost, že vykazují vysokou míru spolehlivosti. Získat je však je často bohužel velmi obtížné. Dosáhnout plné srovnatelnosti experimentální a kontrolní skupiny je téměř nemožné. Experimentální postup je nutné použít již od samého počátku sledovaného programu tak, abychom identifikovali rozdíly mezi těmi, kteří mají z programu užitek, a těmi, kteří ho nemají. U konceptu kontrolní skupiny mohou vzniknout etické problémy. Poskytnout např. sociální výhody jedné skupině a druhé nikoli může být nelegální nebo neetické.

Tyto obtíže vedly k využívání několika skupin kvaziexperimentálních postupů. „Robustnější" kvaziexperimentální postup je založen na srovnání stavu před uskutečněním programu a po jeho realizaci s neekvivalentní (nenáhodnou) kontrolní skupinou. Tento přístup se vzdáleně podobá metodě experimentálního pojetí, kterou jsme popsali výše. Metoda výběru kontrolní skupiny však může být méně precizní a pro vyrovnání počátečních rozdílů mezi oběma skupinami lze využít statistické postupy. K dalším použitelným postupům patří analýza časových řad a porovnání ex post mezi několika skupinami. První z těchto postupů zahrnuje sběr informací a analýzu časových řad, což slouží k tomu, abychom mohli identifikovat změny nebo trendy v chování, které jsou bezprostředním důsledkem zkoumané politiky nebo programu.

c) Kauzální modely

Alternativou k experimentálním modelům popsaným výše jsou kauzální modely. Tyto modely se pokoušejí měřit dopad řady faktorů (nezávisle proměnných) na výsledky určité vládní politiky nebo programu (závisle proměnných). Program sám o sobě je jen jedním z faktorů, které určují výsledek. Podle vhodnosti lze použít některý z řady různých modelů, např. simulační modely, modely *input-output*, mikroekonomické modely a statistické modely. Řady modelů se např. používá nebo používalo k hodnocení dopadů politik strukturálních fondů EU. Jde např. o počítačové modely všeobecné rovnováhy (*CGE – Computable General Equilibrium*), které se v některých případech používají ke stanovení dopadů daňové politiky nebo programů ovlivňujících redistribuci důchodů. Kauzální modely se obvykle využívají v situacích, kdy již lze doložit vztah mezi závisle a nezávisle proměnnými. Používání těchto modelů je však spojeno s rizikem, a proto je na místě obezřetnost. Není-li model dobře specifikován a promyšlen a nevychází-li z řádné analýzy vztahů mezi proměnnými a parametry, mohou být jeho výsledky zavádějící. Kauzální modely by se měly obecně používat jako doplněk k dalším metodám hodnocení.

d) Ekonomické metody hodnocení

Ekonomické hodnocení přináší do metodologie hodnocení informace o nákladech a přínosech. Provádí se buď odděleně, nebo jako doplněk ostatních metod hodnocení. Metody analýzy nákladů a přínosů a nákladové efektivnosti lze tudíž použít ex post, a to k určení toho, zda skutečné přínosy programu opravňují vynaložené náklady.

e) Nekauzální přístupy

V některých případech vzniká problém „zacyklování“, což ztěžuje nalezení přímého vztahu mezi programem a jeho důsledky. Zjištěné výsledky mohou být zčásti důsledkem programu, ale program mohou také ovlivnit externí faktory, které mají na těchto výsledcích určitý podíl. Často např. zjišťujeme, že města, okolo nichž vede dálnice, vykazují rychlejší ekonomický rozvoj než města ostatní. To však může být i důsledkem toho, že trasa dálnice byla navržena tak, aby byla v blízkosti měst, která mají značný potenciál pro ekonomický růst. S problémem pohybu v kruhu se nedokáže vypořádat ani experimentální pojetí, ani kauzální modely.

V takových případech se často uplatní pragmatický přístup. Hodnocení neusiluje o nalezení imaginárních dopadů, ale poskytuje důkladný popis programu a široce využívá pohovorů se všemi zúčastněnými, případové studie, analýzy dokumentů, stanoviska expertů atd. Identifikace chování různých skupin zúčastněných subjektů a vysvětlení toho, jak ovlivnilo procesy programu a výsledky, mohou osvětlit skryté kauzální vztahy a jsou důležité pro úspěšné dosažení cílů politik.

Při hodnocení programů výzkumu a vývoje (R&D) by mělo být zřejmé, že výsledky těchto programů jsou z podstaty věci nejisté – kdyby totiž výsledky byly předem jisté, nepotřebovali bychom výzkum. V této oblasti se často používá nezávislé expertní hodnocení, které může přispět k ex post posouzení toho, zda ocenění rizika ve srovnání s možnými přínosy bylo přiměřené.

Kvalitativní (neboli naturalistické) metody hodnocení vycházejí z principu, že svět je podmíněn sociálně a neustále se mění vzájemným působením jednotlivců. Tento přístup může být vhodný tam, kde hodnotitel není předem vázán k určité sadě hodnot nebo výsledků a je připraven pracovat se všemi zúčastněnými na identifikaci těchto hodnot a relevantních „skutečností“, které nejsou objektivní nebo příčinně předurčené. V tomto přístupu nemůže poskytnout objektivně „správné“ odpovědi, ale může místo toho působit jako zprostředkující mechanizmus k dosažení konsenzu se zúčastněnými subjekty. Ke kvalitativním metodám hodnocení patří pozorování účastníků, etnografické metody, neformální pohovory, případové studie a další. Při aplikacích ve veřejném sektoru však otázku identifikace role různých zájmových skupin účastníků komplikuje význam politické dimenze rozhodování, takže využívání prvků kvalitativního přístupu spolu s dalšími metodami může být přínosnější než studie čistě kvalitativní.[9]

Hodnotící studie v rostoucí míře využívají zjišťování názorů těch, kdo mají z programu prospěch. Patří sem pozorování účastníků, pohovory zaměřené na hodnocení kvalitativních aspektů a s nimi související postupy k odhadování hodnot a preferencí těch, kteří mají z programu prospěch. Těmito postupy můžeme získat informace o mnoha faktorech projevujících se na úrovni domácností a místních společenství, které by nebyly kvantitativními postupy postiženy.[10]

5. Sběr údajů a analýza[11]

Po dosažení souhlasu s obecným přístupem a projektem hodnocení je dalším krokem určení údajů, které potřebujeme k získání nezbytných informací. Údaje jsou skutečnosti a statistická data, které lze pozorovat a zaznamenávat. Rozhodnutí o tom, které údaje jsou nejvýznamnější, nastoluje otázky měření a zjišťování bezprostředních důsledků; to jsme probrali dříve v oddílu A o měření výkonu.

Nemůžeme-li získat relevantní údaje ze sekundárních zdrojů, je nutné shromažďovat údaje primární. Získávání primárních údajů však bude obecně dražší než spoléhání se na data ze sekundárních zdrojů; pokud je to možné, měli bychom se získávání primárních dat vyhnout. Plán na získání primárních údajů obvykle obsahuje výběr způsobu získávání údajů (přirozená pozorování a průzkumy), přípravu nástrojů měření (dotazníky, návod pro pohovory a formuláře pro zápis pozorování) a přípravu plánu výběru vzorků.

Pokud nemůžeme z rozpočtových nebo praktických důvodů použít dostatečně velký vzorek nebo pokud potřebujeme hloubkové údaje, můžeme využít případové studie. Tyto studie umožňují hodnotiteli, aby provedl podrobnou analýzu a získal tak cenné informace a vysvětlující hypotézy pro další analytické práce. Případové studie lze také využít ke zkoumání řady specifických činností nebo projektů, a to tehdy, pokud hodnotitel věří, že takto získá celkové informace o programu. Alternativně se lze k případovým studiím přiklonit také tehdy, když si myslíme, že zahrnují významný vzorek, když porovnáváme fungování nějaké organizace nebo když implementujeme program podle „nejlepší současné praxe".

Musíme stanovit specifické metody analýzy údajů (např. analýzu nákladů a přínosů, vícenásobnou regresi nebo analýzu odchylek), a to v závislosti na druhu požadované analýzy a na dostupnosti údajů. Statistická analýza obsahuje zpracování kvantitativních i kvalitativních údajů k popisu jevů a k vyvození závěrů ze vztahu mezi proměnnými. Nestatistická analýza se většinou provádí pro kvalitativní údaje – např. pro podrobný popis činností nebo procesů nebo pro záznam skupinových diskuzí. Existuje několik druhů nestatistické analýzy (analýza spotřebitelské spokojenosti, induktivní analýza apod.). Nestatistická analýza údajů závisí na profesním úsudku hodnotitele mnohem více než analýzy statistické.

Závěrečné zprávy sloužící k prezentaci výsledků hodnotících studií mají často formu předložení značného množství údajů ve stručné formě. K osvětlení klíčových charakteristik údajů lze použít statistické tabulky, grafy a jednoduché statistické postupy, jako je průměr nebo odchylka.

6. Zprávy o výsledcích hodnocení

Příklady informací, které by měly být součástí zprávy o výsledcích hodnocení, jsou uvedeny v boxu 15.3. Struktura těchto zpráv by však měla odpovídat určitým cílům hodnotící studie a potřebám těch, kterým je zpráva určena: neexistuje jediný univerzálně použitelný model.

7. Metodologické problémy a provádění hodnocení

Metodologické problémy jsou vlastní všem přístupům k hodnocení, jsou však řešitelné, pokud budeme respektovat omezení těchto přístupů a pokud témata správně pojmeme. K tomu potřebujeme specifické znalosti a dovednosti získané buď školením pracovníků, nebo tím, že pro vedení procesu hodnocení smluvně zajistíme experty externí. Tyto problémy však neznamenají, že provádění hodnotících studií nestojí za to. I když hodnocení neposkytne definitivní odpovědi, může do diskuze o navrhování a realizaci vládních politik a programů vnést užitečné informace.

a) Obtíže a nebezpečí

Pro společenské vědy jsou problémy spojené s kauzalitou běžné. Jen málokdy lze získat přesvědčivé důkazy o vztazích mezi příčinou a následkem, protože mít pod kontrolou všechny relevantní proměnné je možné jen zřídkakdy. Experimentální přístup k hodnocení je často obtížný, drahý, trvá dlouho a otázkou je, zda ho lze v praxi vůbec aplikovat. I když experimentální pojetí použijeme, zevšeobecnění výsledků nad rámec podmínek experimentu je často nejisté. Kauzální vztahy mezi programem a pozorovaným výsledkem nelze často – hlavně v důsledku výše diskutovaných obtíží s měřením a stanovením jednoznačných výsledků – nezvratitelně doložit.

Další nesnází je rozhodnutí, zda se zaměříme pouze na oficiálně uznané cíle programu (tj. na ty, které jsou obsahem vládních politik), nebo zda vezmeme v úvahu širší hledisko a budeme se zabývat všemi důsledky programu. Druhý přístup poskytuje souhrnnější obraz o výsledcích programu, je však mnohem složitější a časově náročnější. Stanovení přiměřené doby, po kterou program hodnotíme, je sice obtížné,

Box 15.3 PŘÍKLAD STRUKTURY HODNOTÍCÍ ZPRÁVY

Stručné shrnutí

- Přehled a shrnutí celé zprávy
- Diskuze k silným a slabým místům zvolených hodnotících metod

Úvod

- Popis programu ve smyslu potřeb, cílů, systému dodání atd.
- Souvislosti, ve kterých program působí
- Proč se hodnocení provádí a jeho hlavní otázky
- Popis dalších studií, které byly provedeny

Metodika výzkumu

- Projekt a implementace výzkumu a sběr údajů
- Analýza údajů

Výsledky hodnocení

- Zjištění
- Závěry
- Doporučení

Přílohy

- Kritéria hodnocení
- Doplňující tabulky
- Odkazy a zdroje
- Poznámky k použitým termínům

Zdroj: Komise Evropských společenství (1997).

je ale zásadní; je totiž nutné vyčkat, než se relevantní výsledky plně projeví. Užitečnost informací se však může snížit, změní-li se hodnocený program před dokončením hodnotící zprávy nebo před tím, než lze výsledky hodnocení aplikovat.

Rozhodnutí, zda lze zjištění v rámci hodnocení zevšeobecnit, má význam zvláště tehdy, pokud očekáváme, že hodnocení přispěje k budoucím rozhodnutím o politikách. Podmínky, za nichž se program uskutečnil, však nemusejí být nutně pro budoucnost reprezentativní.

b) Kritéria úspěšnosti hodnocení

Závěry hodnotící studie by měly vycházet z dostatečně širokého rozsahu relevantních otázek. Ten, kdo hodnotí, by měl usilovat o získání co nejpřesnějšího obrazu o daných problémech a měl by je zkoumat tak dlouho, jak to omezené finanční a časové zdroje umožňují. Je však důležité zaměřit se na celou šíři problému. Pokud se šíře hodnocení problému obětuje ve prospěch větší hloubky jeho analýzy, mohou být závěry sice přesné, avšak nebudou pro budoucnost relevantní.

Za předpokladu, že hodnocení má napomoci při rozhodování, musejí kritéria výběru vhodné metody hodnocení zajistit, aby hodnocení poskytovalo užitečné informace. Znamená to porozumět prostředí, v němž se rozhoduje a jemuž jsou závěry nebo výsledky procesu hodnocení určeny.

Při zvažování metody hodnocení je nutné zvážit takové aspekty, jako je užitečnost, finanční možnosti a otázky etiky. Určitý přístup je uplatnitelný v rozsahu, ve kterém ho lze efektivně použít bez nepříznivých důsledků a v rámci daného časového omezení. Naše možnosti jsou dány náklady na implementaci hodnotící studie. Použití metody, která je pro danou situaci nejvhodnější, může být nadměrně nákladné. Při hodnocení je navýsost důležitá objektivnost. Z předložených důkazů a použitých předpokladů by mělo být vždy jasné, z čeho závěry vycházejí. Informace a údaje pro hodnocení by se měly shromažďovat, analyzovat a předkládat tak, aby každý, kdo hodnocení provádí a kdo používá stejné předpoklady, dospěl k obdobným výsledkům. Klient, který provedením hodnotící studie někoho pověří, od něj může požadovat radu nebo doporučení. Za těchto podmínek je důležité udržet odstup mezi objektivními zjištěními studie a doporučeními odvozenými z hodnocení samotného nebo od zdrojů informací, např. opatření politik. Jsou-li závěry nejednoznačné, pak je zvlášť důležité, aby byly jasně stanoveny předpoklady.

Ne každý může mít zájem na plném využití hodnocení. Politici často nejsou ochotni připustit, aby se hodnotily citlivé oblasti, diskutovala zjištění hodnotící studie nebo sami neformulují cíle politiky dost přesně. Manažeři se mohou obávat kritiky své práce. V mnoha zemích je nutné získávat pro hodnocení podporu a za hodnocení bojovat. Provedení hodnotících studií vyžaduje dialog mezi těmi, kteří rozhodují, a účastníky programu. Účastníci programu jsou často těmi, kteří mají zájem na výsledku hodnocení; jsou to např. ti, kteří realizují zkoumaný program. S nimi by mělo být konzultováno jak vymezení problémů, kvůli kterým se program provádí, tak i plán hodnocení, protože jsou to obvykle právě oni, kdo hodnotitelům poskytují údaje, a často hrají důležitou úlohu i při interpretaci výsledků a při realizaci doporučení. Realizátoři programu mají občas pocit, že jejich zájmy jsou hodnocením ohroženy. Pokud se stanou aktivními protivníky, mohou někdy i projekt hodnocení sabotovat.

8. Úloha hodnocení v tranzitivních zemích

Rozvinutí kultury hodnocení vyžaduje čas. Vývoj těchto prací v zemích OECD je nerovnoměrný a hodnocení se neprovádějí systematicky. Tranzitivním zemím se tudíž nedoporučuje, aby si jako bezprostřední cíl stanovily rozvoj uceleného systému hodnocení; tyto země mají jiné aktuální priority, jež souvisejí s oblastmi, jako je vybudování efektivního systému vnější kontroly.

Reformní země obecně potřebují změnit strukturu výdajových programů; k dosažení této změny mohou výrazně přispět informace získané analýzami a hodnotícími studie existujících výdajových programů. V řadě zemí by tudíž mohla být žádoucí příprava hodnotících studií např. v oblasti sociální pomoci, zdravotní péče nebo vzdělání.

Institucionální uspořádání pro zpracování hodnotících studií se v jednotlivých zemích liší. Box 15.4 ukazuje některé příklady tohoto uspořádání v členských zemí EU. V reformních zemích je možné zřídit malý

útvar na centrální úrovni, patrně v rámci ministerstva financí, který by zajišťoval expertizy a metodické vedení pro ostatní ministerstva a pomáhal jim v přípravě hodnocení a při určení kritérií hodnocení pro tyto studie.

Box 15.4 INSTITUCIÁLNÍ STRÁNKA HODNOCENÍ

Francie. Národní výbor hodnocení programů byl založen v roce 1999 a odpovídá za přípravu každoročního hodnocení programů na základě návrhů ministerstev a územní samosprávy. Hodnotící studie financuje Národní fond hodnocení programů; studie jsou publikovány. Kromě činností, které koordinuje Národní výbor, vypracovávají hodnotící studie také ministerstva a sektorové hodnotící komise.

Nizozemí. Rozpočtové odbory všech ministerstev odpovídají za koordinaci programu hodnotících studií a zajišťují, aby byla k dispozici nezbytná pomoc, manuály a výzkumné expertizy. Navrhují hodnotící programy pro jednotlivé projekty, podporují pravidelné hodnocení politik a monitorují kvalitu zpracovávaných analýz a jejich aplikace. Jiné útvary ministerstev poskytují pomoc např. v otázkách řízení lidských zdrojů a organizačního řízení, kontroly a legislativy. Nejvyšší kontrolní úřad zkoumá kvalitu metodiky hodnocení, organizaci a řízení hodnotících studií a vydává o nich zprávy.

Velká Británie. Organizační uspořádání procesu hodnocení je diverzifikované a má „polycentrický“ charakter. Hodnotící studie zpracovává Nejvyšší kontrolní úřad, auditorská komise, ministerstvo financí, další ministerstva, agentury a často i územní samosprávy. Neexistuje zde jedna organizace, která by byla odpovědná za koordinaci této práce, i když ministerstvo financí publikuje dokumenty, které slouží jako vodítka. V některých oblastech je hodnocení na velmi dobré úrovni a zjištění z hodnotících studií se používají při stanovení (a úpravě) priorit politik a při řízení rozpočtu.

Zdroje:
Francie: Conseil Scientifique de l’évaluation (1998).
Nizozemí: OECD (1999).
Velká Británie: Joint seminar of the Tavistock Institute and the Conseil Scientifique de l’évalutation, Paříž, leden 1998.

POZNÁMKY

1. Canadian Institute of Chartered Accountants, 1995.

2. „Hodnocení výdajových programů EU“, Evropská komise, 1997.

3. Tam, kde se vládní programy realizují prostřednictvím sítě srovnatelných institucí v různých regionech nebo na různých územích (např. školy nebo úřady pro poskytování sociálních dávek), lze zavést „interní“ referenční kritéria, to znamená, že školu A můžeme srovnávat se školami B a C. Tento postup se používá např. ve Velké Británii k vytvoření vnitřního trhu, k podpoře konkurence a k vypracování standardů služeb, např. ve zdravotní péči, ve vzdělání, při výběru daní a při výplatách sociálních dávek.

4. Čerpáno z Tilley (1995). Viz také Likierman (1993b).

5. Viz OECD (1999).

6. Např. ve Francii se „hodnocení veřejných politik“ týká jak hodnocení programů, tak hodnocení politik. Conseil Scientifique de l’évaluation (1996).

7. Komise Evropských společenství (1997).

8. Viz Kanada, Ministerstvo veřejných prací a vládních služeb (2000) a Valadez a Bamberger (1994).

9. Viz např. Weiss (1998), zvláště kapitola 11 o „kvalitativních metodách“.

10. Squire (1995) argumentuje tím, že při přípravě hodnocení by se tyto informace musely použít v souvislosti buď s experimentálním, nebo s kvaziexperimentálním hodnocením.

11. Kanada, Ministerstvo veřejných prací a vládních služeb (2000).

POZNÁMKY ČESKÉ REDAKCE

1* NHS – National Health Service – Národní zdravotní služba