

International Price Comparisons for Pharmaceuticals

Measurement and Policy Issues

Patricia M. Danzon¹ and Jeong D. Kim²

1 Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania, USA

2 Yong Sei University, Seoul, Korea

Abstract

Cross-national price comparisons for pharmaceuticals are commonly used for two purposes. Comparisons based on a sample of products are used to draw conclusions about differences in average price levels. Cross-national comparisons applied to individual products are also used by governments to set domestic prices.

This paper examines the major methodological issues raised by international price comparisons, focusing on measurement of differences in average price levels and the validity of policy conclusions drawn from such price comparison studies. It argues that valid measures of average price levels can only be obtained from comparisons based on a comprehensive or representative sample of products, appropriately weighted, following standard index number methods. Comparisons of individual product prices should take into account the manufacturer's entire product portfolio over time rather than focus narrowly on a single product at a point in time.

Because of the great variation across countries in both the range of drug compounds available and the dosage forms, strengths and pack sizes for each compound, obtaining a broadly comprehensive or representative sample is problematical. If products are required to match on all dimensions, including molecule, manufacturer, strength and pack, as is common in most international price comparisons, then only a very small and unrepresentative sample of the drugs available in each country can be included in the analysis. A trade-off between the desire to compare only identical products and the need to compare a truly representative sample of a country's pharmaceutical market is therefore necessary. A valid comparison of average drug prices should include generics and over-the-counter products that are good substitutes for branded prescription drugs, with all forms, strengths and packs. To achieve this broad representation, however, the requirements of same manufacturer, same brand, dosage form, strength and pack size must be dropped.

When such an approach is taken to the comparison of international drug prices, quite different results from those obtained from less comprehensive comparisons may be obtained. Indeed, a major conclusion of this analysis is that international drug price comparisons are extremely sensitive to choices made about certain key methodological issues, such as sample selection, unit of measurement for price and volume, the relative weight given to consumption patterns in the countries being compared, and the use of exchange rates or purchasing power parities for

currency conversion. In particular, the results of this analysis indicate that recent reports suggesting that manufacturer prices in the US are 32% higher than in Canada and 60% higher than in the UK are in fact overstatements which arise from limitations of the sample and methods used to calculate these price differentials.

International price comparisons for pharmaceuticals are used for 2 primary purposes. First, price comparisons based on a sample of products are used to draw conclusions about differences in average price levels, often as input to evaluation of alternative regulatory systems for drug prices. For example, the General Accounting Office (GAO)^[1,2] studies comparing prices in the US with those in Canada and the UK, respectively, were used in the healthcare reform debate in the US. Second, cross-national comparisons applied to individual products are used by governments – for example, in Italy, Spain, The Netherlands and Canada – for setting domestic prices, usually for newly launched products.¹

This paper examines the major methodological issues raised by international price comparisons. Our focus is mainly on the first question, i.e. the measurement of differences in average price levels. We illustrate some of the issues using Intercontinental Medical Statistics (IMS) data for 7 countries and comment on the validity of policy conclusions that are often drawn from such studies. While the cross-national comparison of prices for a single drug for purposes of price regulation appears to be methodologically simpler, it is in fact far more complex when placed in the policy context of setting appropriate prices. We also briefly summarise the normative analysis of determining appropriate price differentials between countries, an issue which is discussed in more detail by Danzon.^[4]

Our main conclusions about broad comparisons of average price levels are based on standard index number theory. To provide a valid measure of

average price levels, the comparison must use a comprehensive or representative sample of products. Instead, most previous studies have focused on small samples of leading, branded, on-patent prescription (Rx) drugs. If individual product prices are to be compared, comparisons based on the manufacturer's entire product portfolio and which take into account lifetime trends are more appropriate than comparisons that focus narrowly on a single product at a point in time. The portfolio approach recognises that many costs of pharmaceutical research and development (R&D) and production are joint costs that cannot be allocated to individual products. Appropriateness of prices depends on the return on capital, which depends on life-cycle revenues for the full portfolio.

Meeting the objective of a comprehensive or broadly representative sample raises particular problems in the case of drugs because of the great variation across countries in the range of compounds available and in the dosage forms, strengths and pack sizes for each compound. As we show, a trade-off is unavoidable. If products are required to match on all dimensions, including molecule, manufacturer, strength and pack, as is common in most previous international price comparisons, then only a very small and unrepresentative sample of the drugs that are available in each country can be included in the analysis. We show that with even the broadest reasonable definition of matching drugs, requiring only that drugs have the same active ingredient and therapeutic category, almost half of the market in France, Germany, Italy and Japan cannot be included in comparisons with the US.

A major conclusion of this analysis is that international drug price comparisons are extremely sensitive to choices made with respect to certain key

¹ President Clinton's Health Security Act (1993)^[3] proposed a similar system, i.e. use of the lowest price in 22 other countries as a benchmark for evaluating the 'reasonableness' of new drug prices in the US.

methodological issues, such as sample selection, unit of measurement for price and volume, the relative weight given to consumption patterns in the countries being compared, and the use of exchange rates or purchasing power parities (PPPs) for currency conversion. Given the inherent cross-national differences in drug markets, there is no single ideal measure of price differences. But while all such comparisons are imperfect, certain methods are more inappropriate than others, depending on the purpose of the comparison.

1. Previous Studies

Index number theory provides widely accepted methods for comparing average price levels in 2 different contexts. These standard methods are used by statistical agencies worldwide to construct their own indexes of change over time in domestic price levels.^[5,6] The same principles apply to comparisons of drug prices across countries, although the implementation is more problematical. However, most previously published studies that attempt to compare average price levels for drugs cross-nationally violate basic index number principles, because of either neglect or lack of data. Most studies have been based on small samples of leading, branded, Rx drugs, with each drug represented by a single pack. By construction, such samples are not representative; in particular, they exclude generic and over-the-counter (OTC) drugs that are important substitutes for branded Rx drugs in some countries. Weighting has been either ignored or inaccurate.

GAO^[1,2] set out to compare manufacturer prices in the US with those in Canada and the UK, respectively, for the leading 200 drugs in the US, by 1991 volume of scripts. Only 121 and 76 drugs for Canada and the UK, respectively, met the matching criteria (which included availability from the same manufacturer in both countries) and hence could be included in the study. For the US, the measure of manufacturer price used was a wholesale list price, which is gross of the discounts and rebates given to managed care, Medicaid and other large purchasers, and hence is upward biased for the net

price to manufacturers. For each drug, the study selected a single dosage form, strength and pack size that was common in the US. This was compared with an imputed price for a similar pack in Canada, based on the Ontario Drug Benefit formulary. Since most of the Ontario formulary prices were per unit (tablet or capsule), a price per pack was imputed by multiplying the per unit price by the number of units per pack. This linear imputation tends to understate prices in Canada, since the formulary price is usually based on the largest pack size which has the lowest unit price.

For the US-Canada comparison, GAO^[1] used 2 measures of average price: (i) the (unweighted) sum of prices in the US relative to the sum of prices in Canada, and (ii) the median of price relatives. The first measure is generally not invariant to the units of measurement, a desirable property of index numbers. Diewert^[5] notes that for this reason, 'virtually no one uses (this) formula at present'. The median of the price relatives is unstable across samples because it focuses on the single product that happens to be the median in the sample selected. The deficiencies of this methodology were implicitly recognised in the UK study (GAO^[2]), which constructed a weighted average using approximate US expenditure weights. Although the expenditure weights pertain to all packs, the price in each country was based on a single pack, so selection bias remains because of use of a nonrandom sample of drugs and probably a nonrepresentative pack for each drug. To illustrate the effect of generics, the UK study substituted one generic for each multisource branded product. The correct approach would include all generics, weighted by market share. The focus on leading products in one country leads to systematic bias, because the identity of leading products differs across countries. These and other issues in previous studies are discussed elsewhere.^[7-9]

The GAO concluded that manufacturer prices in the US are 32% higher than in Canada and 60% higher than in the UK. Our results indicate that these GAO estimates overstate the US price differential,

particularly for Canada, in ways that are not surprising given the limited sample and methods.

2. Principles of Price Comparisons

2.1 Basic Price Indexes

A price index is designed to measure the change in the average price for a market basket of products in 2 situations, the base and the comparison, when prices for most products differ between the 2 situations. An ideal price index measures the change in expenditure required for a consumer to maintain a constant level of utility, when faced with a new price vector. The problem is that when prices change, quantities and utility also change. Unfortunately, utility and the hypothetical expenditure and quantities that would just maintain the base level of utility under the new price vector are not observable. Actual index numbers therefore make use of the observable quantities and expenditures to approximate the ideal index. The widely used Laspeyres index uses the base period/country quantity weights, the Paasche index uses the comparison period/country quantity weights, and the Fisher and Divisia indexes use an average of the observed quantities as weights.

Different indexes can be evaluated according to whether they have certain desirable properties.^[5,10-11] An advantage of the Laspeyres and Paasche indexes is that they bound the unobservable, ideal index. A disadvantage of the Laspeyres and Paasche indexes is that they are intransitive, i.e. the product of the indexes for the US relative to France and for France relative to Germany does not yield the index for the US relative to Germany. Both also fail the country reversal test, i.e. the results depend on which country is used as the base. The Fisher index, which is the geometric mean of the Laspeyres and Paasche indexes, is also intransitive but passes the country reversal test. For example, the index for the US is the reciprocal of the index for Canada: $F_{US} = 1/F_C$.

Despite this theoretical advantage of the Fisher index, the Laspeyres has both practical and theoretical advantages for comparisons of drug prices.

First, it weights the prices of individual products based on quantities of the country selected as the base, for which accurate data are more likely to be available if this is the country undertaking the study. Second, if drug consumption is inelastic with respect to changes in relative prices, because of the role of medical norms, etc., then the Laspeyres index, which uses the base country's quantity weights, may be more appropriate for that country than the Fisher index, which implicitly assumes that consumption patterns would be identical in the 2 countries under comparison if they faced the same prices. Multilateral comparisons raise more complex theoretical problems^[12] and practical problems of noncomparable data. Here we report results for all 3 indexes.

2.2 Sample Selection Bias

2.2.1 Unavoidable Sample Selection Bias

Prices can only be compared for 'international' drugs that are available in both countries under comparison. In practice, a significant fraction of each country's drug expenditure is on 'domestic' products that are used in few, if any, foreign countries and hence cannot be included in cross-national price indexes. Indexes that necessarily exclude a significant fraction of the market in any country will be biased if prices for the omitted, domestic products are systematically different from prices for the included international products, which is likely. Products that diffuse to many countries are disproportionately innovative products,^[13] they are likely to have relatively high prices or volumes, hence may have a significant effect on drug expenditures and be more stringently regulated. The potential bias associated with omitting domestic drugs from the price comparison is therefore probably greatest in countries with regulated prices and a large market share of domestic drugs, such as France and Italy.

The problem of nonmatching products in international price comparisons is sometimes addressed by imputing prices for the missing products. For example, in constructing PPPs, the Organisation for Economic Cooperation and Development (OECD)

bridging methodology applied to multilateral comparisons imputes prices for drugs that are missing based on prices for products that match in at least one bilateral comparison. However, the assumption which underlies bridging methods, i.e. that the missing items would have price relatives similar to those of the matching items, may be reasonable for some industries but is almost certainly incorrect for drugs. It is more likely that certain drugs are unavailable in particular countries because these drugs would not meet that country's regulatory requirements or medical norms. Moreover, to construct weighted price indexes as reported here would also require imputing quantities for missing drugs, which is clearly inappropriate. We therefore do not attempt to impute prices when they are unavailable and do not recommend such imputations, preferring to confine the comparison to those compounds that are observed in both countries under comparison, even if the resulting sample is small.

2.2.2 Avoidable Sample Selection Bias

The unavoidable selection bias that results because some products are unavailable in some countries is exacerbated in many studies by inappropriately restrictive sample selection from the universe of matching products. In particular, most previous comparisons have excluded generics, although generics account for over one-third of Rx sales in some countries and a larger share of scripts. Excluding generics is inappropriate if the aim is to compare the average cost of drug therapy, since generics are chemically close substitutes for the originator brand and indeed are treated as perfect substitutes by third-party payers in many countries. Such exclusion leads to systematic bias in comparisons of average price levels because generics are typically cheaper than their branded counterparts. In particular, exclusion of generics tends to bias upwards the measure of relative prices in countries that have a relatively large generic market share and/or relatively low generic prices, such as the US.^[14,15]

Several studies confine the comparison to leading drugs and/or recently launched drugs. Since the life-cycle price profile for drugs follows a very dif-

ferent path in different countries, this is likely to introduce age-related bias that can severely bias the conclusions about average price levels.^[16] For example, Japan has relatively high launch prices but the post-launch decline in prices in that country is steeper than in most other countries.^[16] Consequently, using a representative sample of drugs with Japan quantity weights leads to the conclusions that Japanese prices are comparable with European prices, whereas a comparison that is restricted to selected, recently launched products might conclude that Japan has relatively high prices.^[17]

2.2.3 Nonmatching Forms and Strength

The ideal unit for comparing prices is a quality-constant course of therapy. In fact, the range and mix of dosage forms, strengths and pack sizes differ significantly across countries. Most existing studies compare the price for a single pack for each product, usually selecting a typical pack in the base country. But this pack may be atypical or totally unavailable in other countries; even if it is available, the distribution of pack sizes and their relative prices may differ, leading to differences in the weighted average price. For example, pack-splitting by pharmacists, who buy large packs and then dispense smaller quantities to individual patients, is common in the US and the UK, but is not permitted in France and Germany, which consequently have smaller average pack size. If the comparison is restricted to a single pack per drug that is identical in all respects in the countries under comparison, the available sample is very small and unrepresentative. Imputation based on price per pill is often used but leads to bias if the relation between unit price and volume is nonlinear because of economies of scale in packaging and/or volume discounts.

2.3 Producer vs Consumer Price Levels

If the policy concern is with the budget impact of pharmaceuticals to payers and consumers, the comparison should be based on retail prices, including payments by payers and any patient co-payment. However, retail prices include whole-

sale and retail distribution margins, which differ significantly across countries and can account for over 50% of the final retail price. Value-added taxes may also be included. Thus, a comparison of retail prices provides at most a very inaccurate comparison of manufacturer price levels.

A comparison of manufacturer prices is of interest if the policy concern is with the contribution of manufacturer prices to the cost of drug therapy to consumers. Given this objective, standard consumer price index methods applied to manufacturer prices provide the appropriate methodology.

3. Data and Methodology

Any international price comparison must define the sample, the unit for measuring price and volume, and the criteria for matching drugs across countries. In all of these choices there is a trade-off: the more precisely the unit of measurement is defined (same compound, same manufacturer, same dosage form, pack size and strength), in order to assure strict comparability of the products across countries, the smaller and less representative is the fraction of each country's market that can be included in the analysis because of differences in product mix. The problem compounds as more countries are included in the comparison. We illustrate these choices and the trade-offs that arise using IMS data.

3.1 Data

The indexes reported here are based on IMS data for sales of cardiovascular products for the year October 1991 to September 1992. Prices are at the manufacturer level, except that for the US these data overstate net prices to manufacturers because they are reported gross of discounts to managed care and other large purchasers, and sales through mail-order, health maintenance organisations (HMOs), supermarkets and other nonpharmacy outlets are omitted. We select only single molecule products. Multiple molecule products are excluded because the relative mix of the active ingredients is not uniform across countries.

Our sample includes generic and some OTC products for both theoretical and practical reasons. Theoretically, it is appropriate to include generics and those OTC products that are close substitutes for originator, Rx drugs. As a practical matter, excluding OTCs would be problematical because a given drug may be Rx in one country and OTC in another. Since OTC status is only reported in our database for two countries, we cannot systematically identify OTCs. However, our sample selection criteria are expected to eliminate most of the OTCs that are more appropriately viewed as 'consumer products' rather than medicines, through the exclusions of multiple molecule products and products that do not match across countries. We also excluded products with sales of less than 1000 packs, less than 1 kilogram of the active ingredient, or missing data on key variables.

3.2 Definition of Product and Matching Criteria

Many molecules are available as different products produced by multiple manufacturers, and most products are available in several different dosage forms, strengths and pack sizes. As shown below, if the unit for comparison is the product, defined by manufacturer, molecule, dosage form and possibly also strength and pack, then only a very small fraction of the drug market in each country is available for comparison.

We use 2 criteria for defining the unit of observation and matching across countries to illustrate the sample attrition that results from narrow product definitions. These are discussed in the following sections.

3.2.1 International Product Name

This criterion defines the product by the IMS international product name (IPN). IMS assigns 2 products the same IPN if 2 of 3 conditions are met: (i) same chemical composition, (ii) same brand name, and (iii) same corporation, including majority-owned subsidiaries. The IPN definition, which matches products across countries only if they have the same manufacturer or brand name, eliminates all licensed products that the licensee sells under a

different brand name from the originator. Also eliminated are all generics, except the unbranded generics with unknown manufacturer, for which IMS assigns the chemical name as the IPN. The IPN definition thus systematically excludes a large and nonrandom subset of products.

3.2.2 Molecule + Therapeutic Category

Our second definition identifies the product by the active ingredient (molecule) and 4-digit anatomical therapeutic class (MOL/ATC). Because the MOL/ATC criterion does not require the same manufacturer or brand in order to match products across countries, all forms of a given molecule including generics and licensed products can be combined to form a weighted average price per MOL/ATC. We required that a molecule be classified into the same therapeutic category to be considered the same product, because market and regulatory factors may differ across therapeutic categories. However, the results are virtually unchanged if we drop the ATC requirement and define the product simply as the molecule.

3.3 Unit of Quantity and Price

The ideal unit would be a quality-constant (equipotent) course of therapy for a given drug, which should be applicable to all dosage forms and strengths. Such ideal units are not observable.

The IMS data offers 3 possible measures that meet the criterion of being available for all dosage forms and strengths: the number of IMS standard units (SU), number of grams (KG) of active ingredient, and number of packs. Of these, we use SU and grams because these most closely approximate the equipotency criterion. IMS defines an 'SU' as one tablet, one capsule, 5 millilitres of a liquid etc, as a rough proxy for a single dose. To the extent that average strength per pill differs systematically across countries, these two measures yield different estimates of price differences, as shown below. There is no strong *a priori* reason for preferring one of these measures over the other. In practice, the price indexes based on price per gram (PKG) are more sensitive to the sample of drugs included, because the distribution of PKG is more highly

skewed, reflecting differences in potency across drugs.

Both the price per standard unit (PSU) and PKG measures permit aggregation over all dosage forms, strengths and packs, thus avoiding the sample attrition and selection bias that arise if comparisons are restricted to a single pack per drug. We also report indexes based only on matching forms/strength to illustrate the effects of more strict comparability. We calculate the weighted average price for a MOL/ATC as total sales for all products in the MOL/ATC divided by total number of SU (for PSU) or total grams of active ingredient (PKG). This approach implicitly ignores real or perceived differences between generically equivalent products with identical active ingredient and therapeutic category, in order to include licensed products and generics that are produced by different companies in different countries.

An alternative, widely used unit of comparison is the defined daily dose (DDD), defined as the number of grams for either a normal dose or a recommended dose. This measure was not available in our data. Because DDDs as commonly defined do not require equipotency or adjustment for differences in duration of treatment, they suffer from limitations (lack of strict comparability) similar to those of the SU and KG measures used here.^[18]

3.4 Indexes

We report indexes for bilateral comparisons of each country relative to the US, to preserve sample size. If the comparison were based on products that are available in all 9 comparison countries, the sample would be restricted to relatively old, global products.

To illustrate the effects of weighting, we report 3 indexes for prices in each country relative to the US. Specifically, the Laspeyres index uses US quantity weights, the Paasche index uses foreign quantity weights, and the Fisher index is the geometric mean of the Laspeyres and the Paasche indexes. The unweighted mean of the price relatives is also reported. Separate indexes are reported for retail pharmacy and hospital sales, because hospital sales are unregulated in some countries that

regulate retail sales and are subject to different competitive pressures.

4. Comparison of Market Characteristics

Table I reports descriptive statistics of the market in each country. The greater number of molecules in Germany, Japan, Italy and France than in the US reflects the large number of domestic molecules in these countries. In the US, the high number of products per molecule reflects primarily generic competitors, whereas comarketing agreements with local companies are probably a more common reason for multiple products per molecule in countries with strict price regulation. Between 50 and 60% of molecules available in the US are also available in all foreign countries except Swe-

den and Switzerland. Similarly, more than two-thirds of the molecules available in Canada, the UK and Sweden are also available in the US. By contrast, fewer than 40% of the molecules available in Germany, France, Italy and Japan are also available in the US. This demonstrates the importance of domestic molecules in these markets.

When the IPN definition is used, matching products account for less than 50% of sales in the US for 5 of the 8 countries. However, using the MOL/ATC definition, more than 90% of the US market is included in all countries except Sweden. The additional 40% (approximately) of the US market that can be included by using MOL/ATC matching represents licensed products and branded generics. The gain in representation from matching on MOL/ATC rather than IPN is less for other

Table I. Effects of alternative matching. Single molecule cardiovascular drugs, 1992 pharmacy

	US	Canada	Germany ^a	France	Italy	Japan	UK	Switzerland	Sweden
Number of products									
Local products	710	157	619	263	365	449	176	160	109
IPN ^b	354	140	535	250	345	437	163	139	99
Molecules	105	66	198	150	154	158	93	97	60
Products per molecule	6.76	2.38	3.13	1.75	2.37	2.84	1.89	1.65	1.98
Percentage of molecules that match									
Molecules available in the US									
% match ^c		50	61	51	54	56	59	43	41
Molecules available in foreign country									
% match ^d		79	32	36	37	37	67	46	72
Percentage of sales on matching products: A. Total product^e									
US sales									
IPN match		73	45	47	53	42	53	48	39
MOL/ATC match		94	96	95	97	95	96	91	76
Foreign sales									
IPN match		81	26	31	34	32	69	57	60
MOL/ATC match		97	61	47	58	57	89	74	92
Percentage of matching sales: B. Dosage forms^f									
US sales									
MOL/ATC match ^g		76	82	74	78	62	67	51	71
Foreign sales									
MOL/ATC match		76	35	34	41	35	60	51	47

a Data are for former West Germany only.

b IPN = International Product Name (IMS) = same chemical and brand or manufacturer.

c The % of molecules available in the US which are also available in the foreign country.

d The % of molecules available in the foreign country which are also available in the US.

e Aggregated over all dosage forms.

f Matched dosage forms.

g MOL/ATC = molecule and therapeutic category.

Table II. Price indexes relative to the US, single molecule cardiovascular drugs. All dosage forms. 1992 pharmacy

Index	US	Canada	Germany	France	Italy	Japan	Switzerland	Sweden	UK
Matched by IPN									
Laspeyres-KG	1.000	1.062	0.816	0.426	0.611	2.789	0.625	0.718	0.720
Laspeyres-SU	1.000	0.969	0.754	0.620	0.865	0.713	0.791	0.653	0.686
Paasche-KG	1.000	1.046	0.542	0.378	0.418	0.987	0.574	0.562	0.606
Paasche-SU	1.000	0.973	0.308	0.447	0.488	0.555	0.688	0.542	0.620
Fisher-KG	1.000	1.054	0.665	0.401	0.505	1.659	0.599	0.635	0.661
Fisher-SU	1.000	0.971	0.482	0.526	0.650	0.629	0.738	0.595	0.652
Unweighted-KG	1.000	1.009	0.888	0.398	0.566	1.157	0.727	2.277	0.791
Unweighted-SU	1.000	0.948	0.824	0.459	0.651	0.498	0.743	0.780	0.707
N ^a	354.000	68.000	41.000	37.000	48.000	33.000	39.000	37.000	68.000
Matched by MOL/ATC									
Laspeyres-KG	1.000	1.166	0.882	0.502	0.704	1.191	0.747	0.871	0.646
Laspeyres-SU	1.000	0.954	0.828	0.659	0.920	0.740	0.885	0.694	0.587
Paasche-KG	1.000	1.016	0.560	0.449	0.504	0.868	0.644	0.677	0.459
Paasche-SU	1.000	0.908	0.587	0.595	0.656	0.620	0.833	0.667	0.463
Fisher-KG	1.000	1.088	0.703	0.475	0.596	1.017	0.693	0.768	0.544
Fisher-SU	1.000	0.931	0.697	0.626	0.777	0.677	0.859	0.680	0.521
Unweighted-KG	1.000	2.166	1.940	0.636	0.961	3.728	1.389	1.184	0.805
Unweighted-SU	1.000	1.627	1.865	0.707	1.080	0.877	1.566	1.027	0.795
N ^a	105.000	52.000	64.000	54.000	57.000	59.000	45.000	43.000	62.000

a N refers to number of distinct dosage forms/strengths, not number of molecules.

IPN = International Product Name (IMS); **KG** = kilogram; **Laspeyres** = US quantity weights; **MOL/ATC** = molecule and therapeutic category; **Paasche** = foreign quantity weights; **SU** = standard units (IMS).

countries than for the US, because of their relatively large market share of domestic molecules. Even with MOL/ATC matching, more than one-third of sales in Germany, Japan, Italy and France represents products that are not marketed in the US. Since these domestic products are necessarily excluded from cross-national indexes, such indexes may not accurately represent the overall price level in countries with a significant market share for domestic products. If the requirement that products have the same dosage form and strength (regardless of pack size), as well as same MOL/ATC, is added, representation of the US market drops to under 80% for all comparisons except Germany, and representation of foreign markets drops below 50% in 5 of the 8 countries.

These data clearly show that differences in product mix and dosage forms in different national markets lead to unavoidable limitations in cross-national price comparisons. Representation can be

significantly increased to include generics and licensed products if the unit of analysis is defined in terms of active ingredient and therapeutic category (MOL/ATC), regardless of manufacturer or brand. Although the implicit assumption that generics and licensed products are perfect substitutes for originator products is only approximately accurate, the trade-off in terms of representativeness and elimination of bias seems worthwhile.

5. Price Indexes

Table II reports the various price indexes for each country, relative to the US, using both PSU and PKG. Although there are 354 IPNs in the US, fewer than 50 of these are available to match in 6 of the 8 foreign countries. Although the reported number of observations appears smaller for the MOL/ATC indexes, in fact a much larger number of products is included, because each MOL/ATC

observation is the weighted average over all products in that molecule/therapeutic category.

5.1 Effects of Weights

The Laspeyres indexes consistently exceed the Paasche indexes (table II). This pattern, namely that a country's price level appears lower if evaluated using its own quantity weights rather than the comparison country's quantities, is observed in studies of other industries and is known as the Gerschenkron effect. In our data, it occurs regardless of the country selected as the base. One contributing factor is the first 'law' of demand, i.e. that countries consume relatively more of the products that are relatively inexpensive in that country. However, in Germany, Japan and Italy, the difference between the Laspeyres-KG and the Paasche-KG indexes of at least 20 percentage points implies very different consumption patterns in these countries compared with the US, presumably reflecting differences in medical norms, preferences, third-party reimbursement, etc., rather than simply response to price differences. To the extent that cross-country consumption differences reflect factors other than price differences, equalising price levels would not equalise consumption patterns. In that case, the effect of a price change for a given country A – achieved, for example, by regulation based on prices in country B – would be more accurately measured by country A's Laspeyres index, which uses its own consumption weights, rather than by either the Paasche or the Fisher indexes, which reflect foreign country B's consumption weights, and hence implicitly assume that country A would adopt country B's consumption patterns if they faced the same prices.

5.2 PSU vs PKG

For a given sample and choice of index (for example, Laspeyres or Paasche), the PSU and PKG measures can differ because of systematic differences in strength (grams per standard unit) between the 2 countries under comparison. The difference is most dramatic for Japan, which appears 37% cheaper than the US when the Fisher-SU index is

used, but 66% more expensive than the US with the Fisher-KG index. This 90 percentage point difference reflects the use of relatively low strength dosing in Japan, such that a given KG of active ingredient requires more units (pills). Two reasons are commonly given for the relatively weak dosing in Japan. First, because of physiological differences, the Japanese are said to require weaker doses to achieve a given therapeutic effect. Second, drug consumption is relatively high in Japan, in part because Japanese physicians profit from dispensing drugs. Weaker doses provide some safeguard against adverse drug interactions.

5.3 Effects of Product Definition and Sample

Table II shows the effect of dropping the requirement that products match on manufacturer or brand name, thereby expanding the sample to include licensed drugs and branded generics. As previously discussed, for each MOL/ATC unit, the price is the volume-weighted average price over all products within the molecule and 4-digit therapeutic category. Thus, the change in number of items in the index (N) understates the gain in number of local products and volume of sales included.

The MOL/ATC indexes in table II are expected to show higher foreign prices relative to the US than the IPN indexes, assuming that generics have relatively low prices and a larger market share in the US than in most other countries, and that the effect of adding branded generics dominates the effect of adding licensed products. This hypothesis is generally confirmed. Increases are larger using the Laspeyres-KG index, which incorporates only the US quantity weights. The relatively small increase for Canada and Germany and the decrease for the UK are not surprising, since these countries also have large generic market shares. Matching on MOL/ATC brings Japan much more into line with other countries and probably reflects the inclusion of licensed products that are excluded under IPN matching.

Table II further illustrates the sensitivity of conclusions to the unit of measurement. In Canada, for example, the Fisher-KG index is 1.088, whereas

Table III. Price indexes relative to the US single molecule cardiovascular drugs. Matching dosage forms only. 1992 pharmacy

Index	US	Canada	Germany	France	Italy	Japan	Switzerland	Sweden	UK
Matched by MOL/ATC									
Laspeyres-KG	1.000	1.067	0.798	0.464	0.735	1.117	0.779	0.719	0.752
Laspeyres-SU	1.000	0.934	0.717	0.640	0.862	0.650	0.883	0.683	0.594
Paasche-KG	1.000	1.003	0.815	0.423	0.741	0.895	0.686	0.685	0.255
Paasche-SU	1.000	0.896	0.715	0.498	0.885	0.469	0.857	0.729	0.456
Fisher-KG	1.000	1.034	0.807	0.443	0.738	1.000	0.731	0.702	0.438
Fisher-SU	1.000	0.915	0.716	0.565	0.873	0.552	0.870	0.706	0.521
Unweighted-KG	1.000	1.446	1.805	0.643	1.144	1.401	0.775	1.386	0.934
Unweighted-SU	1.000	1.616	1.640	0.651	1.120	0.883	1.356	1.116	0.819
N ^a	59.000	59.000	59.000	56.000	48.000	49.000	40.000	39.000	64.000

a N refers to number of distinct dosage forms/strengths, not number of molecules.

KG = kilogram; Laspeyres = US quantity weights; MOL/ATC = molecule and therapeutic category; Paasche = foreign quantity weights; SU = standard units (IMS).

the Fisher-SU index is 0.931. This 16 percentage point gap indicates that cardiovascular drugs are typically weaker in Canada than in the US. Similarly, Japan is 1.7% more expensive than the US based on PKG, but 32.3% cheaper based on PSU, because average strength per dose is lower in Japan.

The indexes that show the largest increases in foreign prices relative to the US, under MOL/ATC rather than IPN matching, are the unweighted price relatives, which now exceed 1.000 for all countries except France, the UK and Italy. These unweighted means are extremely unstable and do not satisfy the country reversal test.

Of our indexes, the Laspeyres-KG indexes with the IPN definition are most similar to those used in GAO,^[1,2] because these IPN indexes require the same manufacturer. Significant differences remain, however. For example, our indexes are confined to cardiovascular products but incorporate all matching products, including generics, within the cardiovascular category. With these caveats, the differences are worth noting. Our indexes imply that Canadian prices are 6.2% higher than US prices, whereas the GAO concluded that US prices are 32% higher than Canada (or Canadian prices are 24% lower than the US). We find that UK prices are 28% lower than the US, whereas the GAO reported that US prices are 60% higher than the UK (or UK prices are 37% lower than the US).

When our preferred MOL/ATC product definition and the Laspeyres-KG index (table II) are used, the foreign price differentials relative to the US are: Japan +19.1, Canada +16.6, Germany -11.8, Sweden -12.9, Switzerland -25.3, Italy -29.6, the UK -35.4 and France -49.8. With the Laspeyres-SU index, all countries are less expensive than the US but the range is smaller, with the UK lowest at -41.3%. These estimates overstate US prices relative to foreign prices because of the omission from the US data of discounts to managed-care and government purchasers and sales through mail-order and HMOs, which are generally at lower prices than retail pharmacy sales.

5.4 Matching on Dosage Form vs Product Totals

Table III reports the MOL/ATC indexes for the sample confined to matching forms/strengths, but retaining aggregation across pack size. Products are thus more closely comparable but there is a significant decline in the number of molecules represented and market coverage (table I). When tables III and II are compared, matching on dosage form generally reduces price differences (the indexes move closer to unity) but not always. For example, the German Fisher-KG index increases from 0.703 to 0.807, and the Italian index from 0.596 to 0.738; the Canadian Fisher-KG index

Table IV. Price indexes relative to the US single molecule cardiovascular drugs. All dosage forms. 1992 hospitals^a

Index	US	Canada	Germany	Italy	Japan	Switzerland	UK
Matched by MOL/ATC							
Laspeyres-KG	1.000	1.242	0.677	0.582	1.389	1.044	0.978
Laspeyres-SU	1.000	0.971	1.014	0.883	0.974	1.409	1.983
Paasche-KG	1.000	0.750	0.543	0.472	1.022	0.903	0.740
Paasche-SU	1.000	0.460	0.554	0.498	0.626	0.871	0.734
Fisher-KG	1.000	0.965	0.607	0.524	1.191	0.971	0.851
Fisher-SU	1.000	0.668	0.749	0.663	0.781	1.108	1.206
Unweighted-KG	1.000	3.866	1.253	0.988	1.664	1.847	1.947
Unweighted-SU	1.000	4.370	1.600	1.222	1.305	2.423	2.573
N ^b	93.000	40.000	44.000	38.000	54.000	23.000	36.000

a Data are not available for hospitals in France and Sweden.

b N refers to number of distinct dosage forms/strengths, not number of molecules.

KG = kilogram; **Laspeyres** = US quantity weights; **MOL/ATC** = molecule and therapeutic category; **Paasche** = foreign quantity weights; **SU** = standard units (IMS).

drops from 1.088 to 1.034 but the Fisher-SU index declines from 0.931 to 0.915.

5.5 Pharmacy vs Hospital

Table IV reports indexes for hospital sales with MOL/ATC matching, which may be compared with the indexes for retail pharmacy sales in table II. The hospital samples are smaller, with fewer matching molecules that meet our minimum volume screens. Conclusions are also more tentative because of greater potential for bias due to discrepancy between list and transactions prices. In the US, discounts are probably more common in the hospital sector than in retail pharmacy sales. Since discounting also occurs in some foreign hospital sales, the direction of bias is less clear than for pharmacy sales, where omission of discounts unambiguously biases upward the measure of US prices relative to foreign prices.

To the extent that generics in the US have a larger share of hospital sales than pharmacy sales, hospital indexes are expected to show higher foreign prices relative to the US. This is confirmed for Japan, Switzerland and the UK. For Canada, the Laspeyres indexes are higher for hospitals than for pharmacies, but the Paasche indexes are significantly lower, particularly on a PSU basis. This suggests that drug consumption in Canadian hospitals is more concentrated on lower priced molecules

and weaker strengths than retail pharmacy sales. The most likely explanation for this greater difference between the hospital and pharmacy indexes for Canada than for other countries appears to be differences in reimbursement incentives; differences in product mix no doubt also play a role, but such differences are likely to occur in all countries.²

6. Comparison of Individual Products

If international price comparisons are used to regulate domestic prices for individual products, the objective of such comparisons should be to establish price differences that are roughly consistent with appropriate price differences in order to achieve welfare-maximising contributions to the joint costs of R&D – so-called Ramsey pricing.^[4] The optimal price differentials depend inversely on the underlying price elasticities of demand for innovative medicines, which presumably depend on income, tastes and other factors. Approaching the problem of international price comparisons from the perspective of optimal price differentials sug-

² Canadian hospitals are reimbursed by fixed annual budgets which create incentives to cut pharmacy costs. Out-patient drugs are reimbursed fee-for-service, and physicians may have greater prescribing freedom and be under less pressure to make price-sensitive choices.

gests certain useful methodological guidelines. In particular:

- Comparisons should be limited to countries that are similar with regard to the factors likely to affect demand elasticity for innovative pharmaceuticals.
- Comparisons should be based on the full portfolio of the manufacturer's products, taking into account life-cycle price profiles. Since many costs are joint costs for a portfolio of products, optimal Ramsey pricing would allocate these joint costs across products differently in different countries, if their relative price elasticities differ across countries. The return on these joint costs depends on life-cycle revenues for the entire portfolio.
- Exchange rates are the relevant basis for currency conversion, because exchange rates determine the innovator firm's actual net revenues from foreign sales in terms of domestic currency, and hence the relative country contributions to financing R&D. Moreover, if foreign prices are converted at PPPs, opportunities for parallel trade occur whenever exchange rates fall relative to PPPs.^[4]

7. What Inferences, if Any, Can be Drawn?

The main conclusion from this analysis is that there is no single, right measure of international price differences for drugs. Results are very sensitive to methodological choices with respect to unit of observation (product or molecule), criteria for matching across countries and implied sample, unit for measuring price (standard unit, gram or some other measure), and weights. None of the available options on any of these issues is ideal. These unavoidable problems arise because of the difference in range of products, dosage forms, strengths and pack sizes of drugs in different countries. This implies a trade-off: if products are required to be identical in all respects in order to be included in the sample, then the potentially matching sample is restricted largely to global, branded products sold under the same

brand name by multinational manufacturers in different countries. However, such products account for only a small fraction of total pharmaceutical sales in all countries.

Although a perfect price comparison is not possible, certain methods are better than others. A valid comparison of average drug prices should include generics and those OTC products that are good substitutes for branded Rx drugs, with all forms, strengths and packs, appropriately weighted by market shares. To achieve this broad representation, requirements of same manufacturer, same brand, dosage form, strength and pack size must be dropped. We show that for cardiovascular drugs examined here, matching on molecule and therapeutic category, and measuring price per dose or per gram of active ingredient, permits over 90% of sales to be included for the US, the UK and Canada, and over two-thirds for most other countries. Requiring the same manufacturer or brand reduces representation to less than 50% of the market for most countries. Relaxing product definitions to include similar forms of the same molecule is consistent with the growing practice of third-party payers in many countries, i.e. reimbursing at a common reference price for all generically equivalent drugs.³

Expanding the comparisons to multilateral comparisons – for example, EU-wide – rather than the bilateral comparisons reported here, would require restricting the comparison to drugs available in all countries, which would further limit the size and representativeness of the sample. Moreover, the issue of weighting, which has been shown to critically affect results even in bilateral comparisons, is more problematical in multilateral comparisons.

For the purposes of drawing policy conclusions, price comparisons raise more questions than they answer. Price indexes are designed to measure average

³ Many state Medicaid and private managed care plans in the US reimburse only for the generic, and require that the consumer pay the difference if they want the brand. Similar provisions apply in the reference price system in Germany, the UK National Health Service and some Canadian provincial health plans.

price differences, not to identify the causes of those price differences. At most, as shown above, a comparison across different indexes can suggest the effects of certain factors, such as consumption patterns (weights), average strength per dose, mix of dosage forms, strengths and pack sizes. To identify the role of costs and other factors, multivariate regression analysis is necessary. For example, Manning^[19] estimates the contribution of liability rules to US-Canada price differences; Danzon and Chao^[15] show the effects of number of competing products and other factors.

Just as regulation is only one factor contributing to price differences, price indexes similarly provide at best a partial measure of one effect of regulation. The effects of price regulation on innovation and productivity are discussed elsewhere.^[20,21] The evidence here does suggest that regulation might affect the market share of domestic products relative to international products. Whereas more than 90% of US retail pharmacy sales are for compounds available in all the other countries in this study, and hence presumably significant compounds, more than 40% of sales in Germany, France, Italy and Japan are for compounds not marketed in the US, which is generally considered to be the most desirable market if regulatory requirements for proof of efficacy can be met. The effect of price regulation on consumption patterns is an important topic for further research.

Acknowledgement

The data acquisition and data analysis reported in this paper were supported by a grant from Pfizer Inc. to the University of Pennsylvania.

References

1. US General Accounting Office. Prescription drugs: companies typically charge more in the United States than in Canada. GAO/HRD-92-110, 1992
2. US General Accounting Office. Prescription drugs: companies typically charge more in the United States than in the United Kingdom. GAO/HEHS-94-29, Jan 12, 1994
3. US Government Printing Office. Health Security Act. HR 3600 IH/S 1757, 1993
4. Danzon PM. The economics of parallel trade. *Pharmacoeconomics* 1998; 13 (3): 293-304
5. Diewert WE. Index numbers. In: Eatwell J, Milgate M, Newman P, editors. *The new Palgrave dictionary of economics*. New York: MacMillan, 1987
6. Berndt E, Griliches Z, Rosett J. Auditing the producer price index – microevidence from prescription pharmaceutical preparations. *J Bus Econ Stat* 1993; 11 (3): 251-64
7. Andersson F. Methodological aspects of international drug price comparisons. *Pharmacoeconomics* 1993; 4 (4): 247-56
8. Danzon PM. The uses and abuses of international price comparisons. In: Helms RB, editor. *Competitive strategies in the pharmaceutical industry*. Washington DC: The American Enterprise Institute Press, 1996
9. Rozek RP. A critique of the GAO report on the differences in prices for prescription drugs between Canada and the United States. *J Res Pharm Econ* 1995; 6 (1): 77-91
10. Diewert WE. The theory of the cost-of-living index and the measurement of welfare change. In: Diewert WE, Montmarquette C, editors. *Price level measurement*. Ottawa: Statistics Canada, 1983: 163-233
11. Diewert WE. The theory of the output price index and the measurement of real output change. In: Diewert WE, Montmarquette C, editors. *Price level measurement*. Ottawa: Statistics Canada, 1983; 1049-113
12. Diewert WE. Microeconomic approaches to the theory of international comparisons. Technical Working Paper No. 53. Cambridge, Mass: National Bureau of Economic Research, 1986
13. Barral E. Twenty years of pharmaceutical research results throughout the world. Paris: Fondation Rhône-Poulenc Santé, 1995
14. Griliches Z, Cockburn I. Generics and new goods in pharmaceutical price indexes. *Am Econ Rev* 1994; 84 (5): 1213-32
15. Danzon PM, Chao LW. Does regulation drive out competition in the pharmaceutical market? Working Paper. Philadelphia: University of Pennsylvania, The Wharton School, 1998
16. Danzon PM, Kim JD. Price indexes for pharmaceuticals: how accurate are international comparisons? Working Paper. Philadelphia: University of Pennsylvania, The Wharton School, 1996
17. Danzon PM. Reference pricing of pharmaceuticals in Japan: lessons from abroad. Working Paper. Philadelphia: University of Pennsylvania, The Wharton School, 1998
18. Danzon PM. Defined daily doses for pharmaceuticals: risks and potential. Working Paper. Philadelphia: University of Pennsylvania, The Wharton School, 1996
19. Manning R. Products liability and prescription drug prices in Canada and the United States. *J Law Econom* 1997 Apr; 40: 203-243
20. Thomas III LG. Industrial policy and international competitiveness in the pharmaceutical industry. In: Helms RB, editor. *Competitive strategies in the pharmaceutical industry*. Washington DC: The American Enterprise Institute Press, 1996
21. Danzon PM, Percy A. The effects of price regulation on productivity in pharmaceuticals. Working Paper. Philadelphia: University of Pennsylvania, The Wharton School. In: *Studies in Income and Wealth*. Chicago: NBER. In Press

Correspondence and reprints: Dr *Patricia M. Danzon*, Wharton School, University of Pennsylvania, 3641 Locust Walk, Room 204, CPC Building, Philadelphia, PA 19104, USA.

