The Cultura di Base project
promoted by Fondazione Compagnia di San Paolo

Impact evaluation

February 2023

Contents

ASVAPP • Association for Public Policy Analysis and Evaluation

Corso Vinzaglio 2 10121 Turin (ITA) Tel +39011533191 www.asvapp.org Email asvapp@pec.it Tax Code 97578830016 VAT No. 08169410019

2

Summary

The Cultural Wellbeing Lab (CWLab) is an initiative of the Fondazione Compagnia di San Paolo, whose aim is to explore the relationship between culture and well-being. With this purpose, it promotes the realisation of culture-base projects that may have a positive impact on people.

The project Cultura di Base, promoted by CWLab, was realised in Turin in 2022, and consisted in temporarily moving medical practices to "architecturally intensive" cultural sites. This report summarises the impact evaluation of the project.

Summary

What is the intervention promoted by Cultura di Base?
Between May 2022 and October 2022, the offices of some general practitioners in Turin were temporarily relocated to several cultural sites. In each site, ad hoc places with "recognised quality in terms of space design and the communication of emotions" were set up. While in the waiting room, and during the visit, patients were exposed to the suggestion induced by the "intense architecture practice" (IAP), possibly changing their feeling and perceptions.

What is the effect of Cultura di Base?
The evaluation estimates the effects of a visit in an IAP on patients. Specifically, we estimate the immediate effects. The evaluation focuses on several feelings and perceptions in the waiting room, on the well-being perceived after the visit, as well as on the perception of the degree of "alliance" with the doctor.

Results suggest that feelings and sensations before and after the visit improved. In particular, the perception of the waiting time definitely improved, and perceived well-being after the visit increased as well. On the other hand, the effects on perceived anxiety, which was in any case very low in general, were negligible, as well as on the ability to get distracted while waiting for the visit. There were no effects on the perceived alliance between doctor and patient. One potential explanation for the last result is that the IAP can more easily influence instant feelings, while the degree of alliance with the doctor is made up of a series of sensations and opinions rooted in pre-existing experiences and perceptions and therefore less susceptible to being influenced at the time by a single and short period spent in the IAP.

Who benefits most from taking part?
The effects are widespread although heterogeneous. The main characterisation concerns age: impact estimates are bigger for older patients.

1. A short description of Cultura di Base

The project, promoted by CWLab in the field "Culture for the humanisation of health care facilities", is focused on the offices of general practitioners in Turin. It consists in temporarily moving medical offices from ordinary practices (hereinafter OPs) to places characterised by "intense architecture" (hereinafter IAPs or "intense architecture practices") or high quality in terms of the design of spaces and communication of emotions. These places *"become the doctor's waiting room. While waiting, the patient is immersed in a kinesthetic experience, benefiting from the cultural content of the places involved in the experiment, which amplifies the effect of the architectural space, reduces the stress of waiting, increasing well-being and psycho-physical comfort"*.

The action implemented by the project *"is based on the idea that a well-designed space can be a vehicle for positive feelings connected with the broader concept of health. Based on this assumption, Cultura di Base intends to move the experience of waiting and the visit from the usual places, making them part of an educational and cultural journey in the architectural space, a contributor to mental and physical well-being. The architecture and cultural content of the selected spaces are offered to patients and doctors involved in the project as a multisensory experience that generates deep emotions and learning and therefore contribute to the treatment path itself."*

In summary, moving GP practices and their waiting rooms to places of culture is intended to influence the experience of the medical visit, triggering potential benefits for both the doctor and the patient and the relationship between them.

The project was carried out by a partnership consisting of Fondazione per l'Architettura (lead organisation), Arteco, Circolo del Design, Order of Doctor-Surgeons and Dentists of the province of Turin, Turin Local Health Board. Realised in Turin between May and October 2022, Cultura di Base involved seven GPs working in five IAPs set up at:
- - the Museo Egizio [Egyptian Museum];
- - the Museo dell'Automobile [Car Museum];
- - the Parco d'Arte Vivente [Living Art Park];
- - the Primo Levi Library;
- - the Polo del '900 Cultural Center.

2. Objectives of the analysis

The impact evaluation is intended to estimate the changes on the direct beneficiaries of the project. More precisely, the target population of interest for this evaluation is composed by patients. We focus on patients of the doctors involved in the initiative who made a visit during the Cultura di Base implementation period.

The objective of CWLab is generally to promote initiatives intended to improve the health and well-being of citizens, these conditions being more specifically defined for each project according to its characteristics. In the case of Cultura di Base, the focus is on the conditions of the visiting patients:

- the possibly negative feelings and perceptions of the patient while waiting (perception of anxiety and time dragging on);
- the perceived state of well-being after the visit;
- the perceived degree of "alliance" with the doctor, i.e. the mutual trust and attention and collaboration in treatment strategies.

The analysis must therefore assess whether the conditions of patients visiting an IAP differ from those that would have been observed if the visit had taken place in an OP. Taking as an example the state of anxiety perceived while waiting, the evaluation of the effects answers the questions: *"What is the level of anxiety perceived while waiting in the IAP?", "What would have been the level of anxiety perceived if the visit had taken place in an OP?"*. The difference between the two conditions estimates the effect of the IAP.

ASVAPP • Association for Public Policy Analysis and Evaluation
Corso Vinzaglio 2 10121 Turin (ITA) Tel +39011533191 www.asvapp.org Email asvapp@pec.it Tax Code 97578830016 VAT No. 08169410019

5

3. Evaluation design and data description

The evaluation is essentially based on the comparison between a "treated group", i.e. patients visiting an IAP, and a "control group", i.e. patients visiting an OP. Both groups of patients answer a questionnaire at the end of the visit. Before describing the evaluation design (in Section 3.2), Section 3.1 provides some additional details on outcomes definition and data collection.

3.1. Data collection

The outcomes of interest are state of anxiety, state of well-being, degree of doctor-patient alliance; they are conditions that can in principle be detected through the perceptions (therefore statements) of patients themselves. In our case, data collection relies on a printed questionnaire that patients are asked to fill in at the end of the visit. The first part of the questionnaire is about socio-demographic characteristics, as well as information on the reasons and frequency of visits. The second part focuses on the outcomes of interest, defined and measured as follows:

- the distraction during waiting time (measured with a 4-point Likert scale);
- the feeling that time dragged on while waiting (measured with a 4-point Likert scale);
- the state of anxiety while waiting (measured on a continuous line and translated on a 0-100 scale);
- the state of well-being perceived at the end of the visit (measured on a continuous line and translated on a 0-100 scale)[1];
- the perceived degree of alliance between doctor and patient (measured with the WAI, Workforce Alliance Inventory, a validated scale based on a battery of 12 questions).

To encourage the participation of as many citizens as possible, the questionnaire was translated in several languages (French, English, Spanish, Romanian, Arabic, Mandarin), simultaneously recording whether the patient was a native speaker of the language used. The project volunteers at the IAPs and the general practitioner staff at the Ops were in charge of the operational management of the survey and involvement of patients (a short handbook was provided).

3.2. Evaluation design

A comparison between two conditions is required: the condition observed on subjects exposed to a particular "treatment" (i.e. after the visit to an IAP), the so-called factual situation, and the condition that would have been observed without treatment (if the visit had taken place in an OP), the counterfactual situation. Since the counterfactual term is unobservable, the challenge is to provide a reliable estimate.

Our evaluation design is based on the use of a control group, i.e. patients who visit an OP instead of an IAP. Our strategy relies on the fact that, during the project, the doctors kept both the IAP and the OP open at the same time, receiving patients in one or the other at different times of the day. The treated group therefore consists of patients who visited an IAP, while the controls are the patients of the same doctors who visited an OP. All patients are interviewed after their first visit: if they have completed the questionnaire on a previous visit, they are not required to do it again.

---

[1] To measure the perceived well-being in the CWLab projects, extensive use was made of the validated PGWBI-S scale (the Short version of the Psychological Personal Well-Being Index, based on a series of 6 questions). A different choice was made in the case of Cultura di Base because, while PGWBI-S measures well-being by asking questions about conditions in the last month, by design this evaluation requires an instant assessment to be made of the condition at the time. The PGWBI-S questions were however included in the questionnaire, but for the sole purpose of giving an indication of the pre-treatment patients' condition.

It is important to stress that the evaluation design allows an immediate effect to be estimated: the patient visits an IAP and, immediately after the visit, answers the questionnaire. It is not therefore possible to estimate the effects over a longer period (which is why the measure of well-being cannot be based on questions relating to a period other than the time of the survey), nor is it possible to estimate the effect of prolonged exposure to an IAP (for example, making more visits to an IAP over a certain period of time).

For organisational and practical reasons, it is impossible to randomly allocate patients to one of the two medical practices, therefore we deal with the risk of "selection bias", i.e. that there are initial differences between the two groups, which means that they are not easily comparable (in other words, if the groups do not have the same initial characteristics, it cannot be assumed that in the absence of treatment they would have shown the same outcomes). The potential initial differences arise from the fact that, without a random selection, visiting an IAP rather than an OP depends fundamentally on the characteristics of the patients. For example, a person with a higher propensity to enjoy culture may be more oriented towards choosing an IAP, while availability in certain time slots may shift the choices of workers or pensioners. It is important to note, however, that i) visits are arranged by booking, and the choice of the practice depends more on time availability in the doctor's and the patient's calendar; ii) the organisation of the IAP and OP calendars ensured, as far as possible, the availability of morning and afternoon slots at both practices. It can therefore be assumed that the initial differences are small, although the need to obtain a solid and credible estimate requires appropriate strategies to be used in order to minimise the selection bias.

In order to produce unbiased impact estimates, we use a series of inference techniques based on the *selection of observables* assumption. In a nutshell, the assumption means that all initial differences between the two groups can be observed. In practice, if exhaustive information on individual characteristics can be observed for each treated and untreated subject, we should be able, when comparing the two groups, to disentangle the effect of the treatment from the effect of initial differences. For the impact estimate we mainly rely on statistical matching, which consists in selecting or weighting the controls in order to create a control group with the same initial characteristics as the treated patients. The similarity between the two groups ensures the comparability of the outcomes, and therefore the possibility of estimating the effects of the project. The credibility of the estimates requires the *selection of observables* assumption to be valid, in order to ensure that all the individual characteristics playing a role in the selection process (so-called control variables) can be observed. In other words, the set of individual information needs to be rich and exhaustive. In our case, the assumption is supported by the possibility of collecting data with an ad hoc questionnaire that collects all the information deemed of interest. They include:

- the main demographic and social characteristics: gender, age, nationality, educational qualification, marital status, employment status;
- the doctor (to check the influence of each individual doctor on the perceptions of patients);
- the relationship with the doctor and the frequency of visits.

Section 4 presents the project impact estimates based on a specific matching technique, the kernel matching. Please refer to Appendix A1 for more technical details and a short discussion of other options. Appendix A1 presents a robustness analysis by comparing the results obtained with different techniques, highlighting the substantial homogeneity of the results.

3.3. Sample description

The survey covered a period of around two thirds of the six months of the trial (the first month of settling-in and the month of August are excluded). The number of surveys completed by the two groups was 617 and

ASVAPP • Association for Public Policy Analysis and Evaluation
Corso Vinzaglio 2 10121 Turin (ITA) Tel +39011533191 www.asvapp.org Email asvapp@pec.it Tax Code 97578830016 VAT No. 08169410019

7

277 respectively. Having booked their visit, patients remained in the waiting room for a very short period: no more than 10 minutes in two thirds of cases, half an hour or more for less than 10%.

Before the analysis, questionnaires filled in by non-native speakers (43, largely treated) and those who refused to fill in the questionnaire (2.6% of the treated and 2.4% of controls, 21 in total) were excluded. Finally, the sample for the analysis has 564 treated and 266 controls. Tables 1 and 2 summarise their main characteristics[2].

Table 1. Demographic and social characteristics

	OP (controls)	IAP (treated)	
Female	58.5%	56.4%	
Male	41.5%	43.7%	
Italian	92.2%	91.5%	
Foreign born	7.8%	8.5%	
40 years old or younger	20.5%	21.4%	
41-60 years old	32.7%	35.6%	
61 years old or older	46.8%	43.0%	
Education: Compulsory school or less	21.8%	27.3%	*
Education: High school	42.9%	38.3%	
Education: Degree	35.3%	34.4%	
Single/widowed	49.2%	45.4%	
Married/cohabiting	50.8%	54.6%	
Not employed/retired	52.5%	51.7%	
Employed	47.5%	48.3%	
No.	266	564	

Note: significant differences for *α=10% **α=5%

Table 2. Initial conditions and relationship with doctor

	OP (controls)	IAP (treated)
Frequency of visits: regular (at least every 2 months)	30.0%	27.0%
Frequency of visits: occasional	70.0%	73.0%
Doctor chosen based on advice or friendship/trust	68.1%	73.3%
Doctor chosen based on convenience or other reason	31.9%	26.7%
State of severe distress in the past month (PGWBI-S)[3]	34.9%	32.4%
State of distress in the past month (PGWBI-S)	22.6%	20.5%
State of no distress in the past month (PGWBI-S)	26.0%	29.9%
State of well-being in the past month (PGWBI-S)	16.6%	17.2%
No.	266	564

Note: significant differences for *α=10% **α=5%

More than 90% of the participants were Italian, the gender distribution was fairly equal (57% women), and the age distribution was fairly uniform, with a predictably higher number of the over 60s. In about 70% of cases, patients had an occasional visit.

---

[2] The questionnaire was planned to involve all patients in the survey, keeping track of all rejections. Even in the absence of total attendance data, it is likely that during this period the total number of patients was higher than that of respondents, and that tracked refusals were only a part of those who did not take part. Furthermore, the difference between the number of treated and controls suggests a greater effort made by the volunteers in the IAP compared to the secretarial staff in the OP in asking patients to complete the questionnaire. Therefore the results of the analysis can be generalised to the population represented by the respondents, in particular the treated ones.

[3] The PGWBI-S indicator assigns values between 0 and 110. Values below 60 indicate severe distress; values between 60 and 69 indicate distress; values between 70 and 89 indicate no distress; values of 90 and above indicate a positive state of well-being.

ASVAPP • Association for Public Policy Analysis and Evaluation
Corso Vinzaglio 2 10121 Turin (ITA) Tel +39011533191 www.asvapp.org Email asvapp@pec.it Tax Code 97578830016 VAT No. 08169410019

8

The characteristics of the sample provide a description of the target population and are useful to check for initial differences between treated (IAP) and controls (OP). These are generally small and, except in some cases, they are not statistically significant. The similarity supports the assumption that the booking procedure, although not a randomisation, induced choices guided above all by logistical aspects, producing two quite similar (and comparable) groups. The matching techniques will then be used to eliminate residual differences and make a comparison between groups under the same initial conditions.

A final note concerns the number of observations for each IAP and for each OP, which is quite heterogeneous (Table 3). Since we estimate the overall average effect of the project, our results will be more influenced by IAPs with more observations. We do not provide separate estimates for each IAP for at least two reasons: i) the sample sizes would be too small; ii) they would be poorly relevant from a policy perspective: the results would depend not only on specific characteristics of the single IAP, which are difficult to codify, but also on those of the respective OP.

Table 3. Number of observations per cultural site

	OP (controls)	IAP (treated)	Total
Primo Levi Library	13	107	120
Museo dell'automobile	40	109	149
Museo Egizio	48	165	213
PAV	30	57	87
Polo del 900	135	126	261
Total	266	564	830

ASVAPP • Association for Public Policy Analysis and Evaluation
Corso Vinzaglio 2 10121 Turin (ITA) Tel +39011533191 www.asvapp.org Email asvapp@pec.it Tax Code 97578830016 VAT No. 08169410019

9

4. Results

The impact estimate is therefore based on a comparison between treated patients and controls, as defined above. The estimates here summarised are obtained by refining the comparison using kernel matching, as explained more fully in Appendix A1. Participants' average outcomes represent the factual condition, while matched controls' outcomes estimate the counterfactual condition: what would have happened to the former if instead of in an IAP they had visited their doctor at an OP. Tables 4 and 5 summarise the characteristics of the two groups after matching. The differences between the groups are smaller than the initial ones. The number of people in the two post-matching groups is smaller. This is because the initial number included cases for which some data needed for the matching analysis were missing. The matched groups show the same composition of the initial ones, therefore the restriction does not affect the estimates generalisability.

Table 4. Post-matching demographic and social characteristics

| | Pre-matching | | | Post-matching | |
	OP (controls)	IAP (treated)		OP (controls)	IAP (treated)
Female	58.5%	56.4%		57.3%	57.4%
Male	41.5%	43.7%		42.7%	42.6%
Italian	92.2%	91.5%		91.6%	90.6%
Foreign born	7.8%	8.5%		8.4%	9.4%
40 years old or younger	20.5%	21.4%		20.9%	20.9%
41-60 years old	32.7%	35.6%		35.2%	36.6%
61 years old or older	46.8%	43.0%		43.9%	42.5%
Education: Compulsory school or less	21.8%	27.3% *		25.6%	26.0%
Education: High school	42.9%	38.3%		38.0%	40.7%
Education: Degree	35.3%	34.4%		36.4%	33.3%
Single/widowed	49.2%	45.4%		46.6%	44.8%
Married/cohabiting	50.8%	54.6%		53.4%	55.2%
Not employed/retired	52.5%	51.7%		49.0%	50.4%
Employed	47.5%	48.3%		51.0%	49.6%
No.	266	564		211	478

Note: significant differences for *α=10% **α=5%

Table 5. Post-matching initial conditions and relationship with doctor

| | Pre-matching | | | Post-matching | |
	OP (controls)	IAP (treated)		OP (controls)	IAP (treated)
Frequency of visits: regular (at least every 2 months)	30.0%	27.0%		27.2%	27.2%
Frequency of visits: occasional	70.0%	73.0%		76.3%	73.6%
The Doctor chose based on advice or friendship/trust	68.1%	73.3%		72.6%	74.3%
The Doctor chose based on convenience or other reason	31.9%	26.7%		27.4%	25.7%
State of severe distress in the past month (PGWBI-S)	34.9%	32.4%		33.5%	33.3%
State of distress in the past month (PGWBI-S)	22.6%	20.5%		22.1%	21.1%
State of no distress in the past month (PGWBI-S)	26.0%	29.9%		29.0%	28.8%
State of well-being in the past month (PGWBI-S)	16.6%	17.2%		15.4%	16.8%
No.	266	564		211	478

Note: significant differences for *α=10% **α=5%

ASVAPP • Association for Public Policy Analysis and Evaluation

Corso Vinzaglio 2 10121 Turin (ITA) Tel +39011533191 www.asvapp.org Email asvapp@pec.it Tax Code 97578830016 VAT No. 08169410019

10

Table 6 summarises the estimates. As regards the distraction in the waiting room, 71% of respondents said they managed to get distracted (either "enough" or "a lot"). We estimate an increase of 4.5 percentage points in "a lot" answers, and a similar reduction in "enough" answers, but these estimates are statistically not significant. The improvement in the feeling that time dragged on is decidedly more evident: those who answer "not at all" increased by 13.5 percentage points (+26%), against an almost equal reduction in the more moderate answers. As regards the state of anxiety, measured on a scale from 0 to 100, the declared level is on average low (10.5 out of 100), and IAPs effects are negligible.

The next outcome is the well-being felt at the time of completing the questionnaire, therefore after completing the visit. The average level, measured on a scale of 0 to 100, is 65 points, with a significant increase of 3 points due to IAPs.

The last outcome is the doctor-patient alliance, detected with the Workforce Alliance Inventory and reported here on a scale of 0-100. The average level measured on this scale is high, more than 80 points. In this case, no significant effects of IAPs can be detected.

Table 6. IAPs impact estimates

		Average level for treated patients	Effect estimate	
Were you able to get distracted while you were waiting for the visit today?	*Not at all*	10.7%	-0.6%	
	A little	18.7%	+0.2%	
	Enough	47.7%	-4.1%	
	A lot	23.0%	+4.5%	
Did you feel that time seemed to drag on?	*Not at all*	65.4%	+13.5%	**
	A little	24.7%	-11.9%	**
	Enough	8.8%	+2.5%	
	A lot	1.1%	-4.1%	
Anxiety perceived while waiting (0-100)		10.5	-0.1	
Well-being perceived after the visit (0-100)		67.4	+3.2	**
Perceived doctor-patient alliance (WAI, 0-100)		83.1	+0.8	
	No.	478		

Note: significant estimates for $*\alpha=10\%$ $**\alpha=5\%$

Summing up, it can be said that the patients experienced fairly positive feelings. They had a better time perception while waiting and showed a higher perceived wellbeing after the visit. No changes in perceived alliance with the doctor were estimated. A possible reason for the last result, which partially contrasts with the previous ones, may be that this analysis can catch the immediate effect of a single visit. The former outcomes are more subject to short-time changes, and therefore they are more easily influenced by a visit, while the perceived alliance, which depends more on a deep-rooted situation linked to respect, trust, pre-existing opinions about one's doctor, is perhaps less modifiable by such a treatment.

4.1. Are the results affected by patient characteristics?

Having shown the average effect estimates of IAPs on patients, this section investigates a more detailed question, i.e. whether and to what extent the effects are heterogeneous among different types of patients. Our estimates were then replicated for subgroups of people. For the sake of brevity, and without loss of generality, the estimates presented in next tables are restricted to a small number of outcomes. In particular, for the questions concerning distraction and time dragging, we focus on answers "a lot" and "not at all" respectively, i.e. the ones on which a more relevant impact was estimated.

Table 7. IAPs impact estimates, by demographic and social characteristics

	Distraction (A lot)	Time dragging (Not at all)	Anxiety (0-100)	Well-being (0-100)	WAI (0-100)	N^T
Female	7.9%*	12.0%*	0.0	5.2*	-0.4	274
Male	-0.1%	15.7%**	-0.4	0.6	2.6	203
40 years old or younger	-5.7%	6.6%	5.1	1.8	-0.3	98
41-60 years old	5.6%	13.9%*	1.1	3.8	2.2	175
61 years old or older	9.3%*	15.8%**	-3.8*	3.7*	0.1	203
Education: Compulsory school or less	14.0%**	10.9%*	-0.7	3.2	-1.4	112
Education: High school	2.0%	4.2%	-1.1	4.7	3.3	192
Education: Degree	1.1%	28.3%**	1.2	2.5	-1.3	174
Single/widowed	-4.2%	9.5%	1.3	3.3	0.5	264
Married/cohabiting	12.2%**	17.3%**	-1.2	3.3	0.4	213
Not employed/retired	2.3%	10.1%*	-0.3	3.8	-2.3	240
Employed	7.2%	15.8%*	0.2	2.8	2.2	237

Note: significant estimates for *α=10% **α=5%

N^T: number of treated patients included in the estimate. The first two columns refer to the "A lot" answer to the question "Were you able to get distracted while you were waiting for the visit today?" and to the "Not at all" answer to the question "Did you feel that time seemed to drag on?".

Table 8. IAPs impact estimates, by initial conditions and relationship with the doctor

	Distraction (A lot)	Time dragging (Not at all)	Anxiety (0-100)	Subjective (0-100)	WAI (0-100)	N^T
Make regular visits	6.2%	8.9%	-3.1	3.4	3.5	130
Have visited occasionally	3.9%	15.6%**	0.8	3.4*	-0.6	348
Doctor chosen based on advice or friendship	7.5%*	8.5%	1.0	2.3	1.0	355
Doctor chosen for convenience	4.7%	27.1%**	-2.5	4.5	-0.1	122
State of distress in the past month (PGWBI-S)[4]	6.4%	9.9%	1.9	3.0	-1.0	253
State of no distress in the past month (PGWBI-S)	3.2%	17.9%**	-2.1	2.0	2.2	211

Note: significant estimates for *α=10% **α=5%

N^T: number of treated patients included in the estimate. The first two columns refer to the "A lot" answer to the question "Were you able to get distracted while you were waiting for the visit today?" and to the "Not at all" answer to the question "Did you feel that time seemed to drag on?".

In general, we can assert that the positive impact of the IAP extends to a wide range of patients, rather than being limited to specific categories. Nonetheless some peculiarities emerge. One is that the older the patient the greater the effect. Older patients are more numerous, therefore the significance of estimates increases, but the effects size increases with age as well. The over-60s show a more evident reduction in anxiety, and when age increases we estimate a higher reduction in the perception of time dragging. Furthermore, the over-60s are the only category that shows a strong improvement in the ability to distract themselves while waiting, and a significant effect on perceived well-being. These differences presumably also explain the differences according to marital status (the effects are higher for married or cohabiting people, over-represented among the over 40s), gender (women under 40 are under-represented compared to men), the previous state of well-being (the "no distress" condition is slightly more frequent among the less young).

The interaction effects of education, positive for those with a low qualification and those with a degree, are not clearly interpretable. One hypothesis is that concomitant and opposing factors may affect them: on the one hand a higher sensitivity of more educated people, on the other the higher incidence of low education levels among older people.

---

[4] The classification adopted here distinguishes only two categories: people with distress or severe distress (PGWBI-S score less than 70/110), people in a state of no distress or positive well-being (score higher than or equal to 70/110).

Finally, better results emerge for those having a visit occasionally. This result would contrast with those relating to age (older people make regular visits more often); actually, in most cases the difference concern more the significance than the effect size. Looking at numbers, we cannot exclude that this is simply a question of significance due to the different sample sizes.

Appendix A1. Details on estimation techniques

The evaluation is based on the use of a control group, i.e. a group of patients who did not receive the treatment. The idea is to use the outcomes of this control group to estimate what would have happened to the treated in the absence of treatment. This appendix presents some operational details that have been excluded from the text for brevity. They relate in particular to the propensity score matching method used to produce the estimates.

In comparing treated to controls to estimate the effects of an intervention, it is important to check that the two groups were similar before the intervention. Then we can assume that without intervention their subsequent outcomes would also have been similar. If initial differences exist, a simple ex-post comparison between groups would lead to estimates affected by selection bias. In our case, with no random assignment of patients to OPs or IAPs, we cannot take the similarity for granted, although some favourable conditions:

- the choice of IAP or OP is presumably based on calendar availability, therefore not depending on patients' feelings and sensitivity;
- the availability of IAPs and OPs at different times should avoid self-selection induced by patients' work conditions.

Our estimates were obtained with techniques based on the selection of observables: assuming that the initial information available on each person can explain all the differences between treated patients and controls, then the information (the so-called control variables) can be used to "clean up" the final comparison.

The technique widely used in this analysis is matching. It is a procedure based on matching each treated patient to one or more controls who are as similar as possible. The matching procedure produce a subset of controls, selected so to have the same composition of the treated patients. The outcomes of treated group and matched control group can then be compared, and the difference can be more plausibly attributable to the treatment.

The matching ensures an unbiased estimate if the selection of observables assumption holds. It is realistically impossible to observe all people characteristics: in practice what can be done is to observe a sufficiently rich and informative set of initial characteristics. In our case we actually have a rich and exhaustive set of information, since the data collection is based on an ad hoc questionnaire (see Appendix A2). We were able to include all the questions considered useful for determining the initial condition of the people involved:

- the main demographic and social characteristics: gender, age, nationality, educational qualification, marital status, employment status;
- the doctor's name;
- the relationship with the doctor and the frequency of visits;
- the date of the visit and the reason for choosing the practice.

The matching strategy used here is based on propensity score. The *pscore* is the method used to "measure" the initial similarity between patients. In practice, the *pscore* is a single variable that summarises all the initial characteristics. Following statistical jargon, the propensity score is the *probability, conditioned on initial characteristics, of being treated*; leaving aside the technical details, the *pscore* can be considered as a summary index, which varies between 0 and 1, constructed in such a way as to appropriately summarise all the information available on a single patient. In particular, the weighting of each individual characteristic in the *pscore* depends on the extent to which that characteristic discriminates between treated patients and controls. Once the *pscore* has been estimated, the difference between two individuals is measured as

ASVAPP • Association for Public Policy Analysis and Evaluation
Corso Vinzaglio 2 10121 Turin (ITA) Tel +39011533191 www.asvapp.org Email asvapp@pec.it Tax Code 97578830016 VAT No. 08169410019

14

difference between their *pscores*. Under suitable (verifiable) conditions, two individuals with a similar *pscore* are on average similar in all characteristics.

In order to estimate the effects via matching, it is then sufficient to match each treated patient to controls having a similar *pscore*, i.e. who have on average the same initial characteristics, whose outcomes represent what would have happened to the treated patient without treatment. The results presented in Section 4 are based in particular on kernel matching: each treated patient is compared to a weighted average of the controls, where the weights increase when the similarity between treated and control increases.

A1.1. Robustness check: do results depend on the estimation technique?

The matching method, widely used in evaluation designs with non-experimental control group, is actually a family of techniques based on the same idea of matching treated patients to similar controls. Distance measures different from *pscore* can be used (even though its use is very widespread). Moreover, many different matching procedures are available. Kernel matching, selected for the final results presented in Section 4, is an example of a specific technique, and requires each treated patient to be matched to a weighted average of controls. Some frequently used alternatives include:
-   nearest neighbour matching, which consists of matching each treated patient to the more similar control (i.e. with the most similar *pscore*);
-   radius matching, which consists of comparing each treated patient to controls within a certain radius, i.e. whose distance in terms of *pscore* is lower than a fixed threshold.

The estimation strategies based on the selection of observables also include the use of traditional regression models. Even though the main assumption on which they are based is the same, they are based on a quite different technique: the regression model can isolate the net influence of every control variable on the outcome (for example the effect of age on well-being), therefore the effect of the treatment on the outcome can be disentangled from the influence of other factors (for example, the comparison between treated and controls does not depend anymore on age differences, because the model has isolated and removed the impact of age on outcomes).

Regression models offer a more traditional approach, and have some strengths and weaknesses; the various matching techniques, more flexible in some aspects, also have strengths and weaknesses depending on what they are used for. Although kernel matching was selected as the "preferred" strategy in our presentation, the analyses were also conducted using other techniques. The objective is to verify the robustness of the estimates, i.e. to check whether they depend on the used technique. As previously stated, the results are substantially the same, whatever the technique, as shown in Table APP1.

Table APP1. IAPs impact estimates, by estimation technique

	Distraction (A lot)	Time dragging (Not at all)	Anxiety (0-100)	Subjective (0-100)	WAI (0-100)
Radius matching (r=0.01)	+1.7%	+17.2%**	-0.2	+4.1*	+0.3
Nearest neighbour matching	+3.4%	+19.2%**	-1.9	+7.3**	+1.5
Kernel matching	+4.5%	+13.5%**	-0.1	+3.2**	+0.8
Simple linear regression	+5.0%	+13.7%**	+0.4	+4.6**	+1.7

Note: significant estimates for *α=10% **α=5%

N^T: number of treated patients included in the estimate. The first two columns refer to the "A lot" answer to the question "Were you able to get distracted while you were waiting for the visit today?" and to the "Not at all" answer to the question "Did you feel that time seemed to drag on?".

ASVAPP • Association for Public Policy Analysis and Evaluation

Corso Vinzaglio 2 10121 Turin (ITA) Tel +39011533191 www.asvapp.org Email asvapp@pec.it Tax Code 97578830016 VAT No. 08169410019

15

A1.2. Do the results change if we focus on the "non-curious"?

All the estimated presented so far involved all visiting patients who completed the questionnaire. The assumption underlying the estimate is that the individual information collected with the questionnaire is sufficient to explain the initial differences between treated patients and controls, in particular the propensity to choose an IAP or an OP. In the questionnaire, a previously unused question was asked which specifically investigates the reason for choosing the doctor's office. Around one fifth of the respondents stated that they selected an IAP for curiosity, the others mainly for logistical reasons.

The level of curiosity is used as a control variable in the matching procedure, therefore reducing the impact of curiosity on estimates. In any case, this section presents the results of the analysis restricted to the "non-curious" only[5]. The idea is that this restriction should eliminate any doubt about any residual influence of curiosity on estimates. Firstly, Tables APP2 and APP3 show the substantial similarity between the "non-curious" and the total population.

Table APP2. Characteristics of the "non-curious"

	All			"Non-curious"		
	OP	IAP		OP	IAP	
	(controls)	(treated)		(controls)	(treated)	
Female	58.5%	56.4%		58.5%	56.1%	
Male	41.5%	43.7%		41.5%	43.9%	
Italian	92.2%	91.5%		92.9%	90.4%	
Foreign born	7.8%	8.5%		7.1%	9.6%	
40 years old or younger	20.5%	21.4%		20.9%	22.3%	
41-60 years old	32.7%	35.6%		32.0%	34.1%	
61 years old or older	46.8%	43.0%		47.1%	43.6%	
Education: Compulsory school or less	21.8%	27.3%	*	22.7%	30.6%	**
Education: High school	42.9%	38.3%		41.3%	37.7%	
Education: Degree	35.3%	34.4%		36.0%	31.7%	
Single/widowed	49.2%	45.4%		49.0%	48.1%	
Married/cohabiting	50.8%	54.6%		51.0%	51.9%	
Not employed/retired	47.5%	48.3%		46.3%	46.5%	
Employed	52.5%	51.7%		53.7%	53.5%	
No.	266	564		247	448	

Note: significant differences for *α=10% **α=5%

---

[5] In the sample analysed so far, 23% of treated patients stated that they made the choice out of curiosity, but so di 8% of controls. Given the poor interpretability of this answer in respect of the latter, even the so-called "curious" among the controls were excluded here.

Table APP3. Characteristics of the "non-curious"

	All		"Non-curious"	
	OP (controls)	IAP (treated)	OP (controls)	IAP (treated)
Frequency of visits: regular (at least every 2 months)	30.0%	27.0%	29.5%	27.0%
Frequency of visits: occasional	70.0%	73.0%	70.5%	73.0%
Doctor chosen based on advice or friendship/trust	68.1%	73.3%	67.2%	70.6%
Doctor chosen based on convenience or other reason	31.9%	26.7%	32.8%	29.4%
State of severe distress in the past month (PGWBI-S)[6]	34.9%	32.4%	35.9%	31.6%
State of distress in the past month (PGWBI-S)	22.6%	20.5%	21.2%	20.7%
State of no distress in the past month (PGWBI-S)	26.0%	29.9%	26.3%	31.6%
State of well-being in the past month (PGWBI-S)	16.6%	17.2%	16.5%	16.1%
No.	266	564	247	448

Note: significant differences for *α=10% **α=5%

Finally, Table APP4 summarises the estimated effects for the "non-curious". Apart small changes in numbers, our results are completely consistent with those presented in previous tables, therefore strengthening our results: the visible effect on the perception of time and the increase in perceived well-being are confirmed, while the changes on distraction were not significant (albeit positive) and those for anxiety and for the perceived alliance with the doctor were close to zero.

Table APP4. IAPs impact estimates for the "non-curious"

	Distraction (a lot)	Time (not at all)	Anxiety (0-100)	Subjective (0-100)	WAI (0-100)
Radius matching (r=0.01)	+4.9%*	+15.2%**	-3.3*	+5.4**	+0.3
Nearest neighbour matching	+5.6%	+18.5%**	-4.2	+7.7**	+0.4
Kernel matching	+4.4%	+16.4%**	-2.5	+5.3**	+1.1
Simple linear regression	+4.4%	+14.9%**	0.0	+5.7**	+1.8

Note: significant estimates for *α=10% **α=5%

N^T: number of treated patients included in the estimate. The first two columns refer to the "A lot" answer to the question "Were you able to get distracted while you were waiting for the visit today?" and to the "Not at all" answer to the question "Did you feel that time seemed to drag on?".

---

[6] The PGWBI-S indicator assigns values between 0 and 110. Values below 60 indicate severe distress; values between 60 and 69 indicate distress; values between 70 and 89 indicate no distress; values of 90 and above indicate a positive state of well-being.

Appendix A2. The Questionnaire

Part pre-completed by surveyor				
Doctor's name		Time:	morning	afternoon
Date		Mother tongue	YES	NO
Place	IAP	OP		
	(specify address - street or square - if the doctor operates in multiple practices)			

Dear patient, we would like to hear your opinion about your experience during the visit.
Please answer all questions freely and honestly.
We inform you that the questionnaire is anonymous.

D01	**Sex**	*F*	*M*			

D02	**Age**	*18-30*	*31-40*	*41-50*	*51-60*	*61-70*	*71 or older*

D03	**Nationality**	*Italy*	*EU*	*Non-EU*

D04	**Marital status**	*Unmarried*	*Married/cohabitant*	*Widow/widower*	*Other*

D05	**Education**	*Compulsory school or less*	*Secondary school certificate*	*Degree*

D06	**Do you have a job?**	*Yes*	*No*	*Retired*

D07. What was the reason of your doctor's appointment today?
- □ *Actions/services related to Covid19*
- □ *A prescription*
- □ *Medical examination*
- □ *To get advice/information*
- □ *To submit the outcome of tests*
- □ *Other*

D08. How often do you go to the doctor (aside from prescriptions)?
- □ *Every week*
- □ *Every month*
- □ *Every two months*
- □ *Occasionally*

D09. Why have you chosen this family doctor?
- □ *Proximity/convenience*
- □ *Recommended by family members, friends, acquaintances*
- □ *Friendship/trust*
- □ *Other*

D10. Your doctor has two offices, why have you come to this office?
- □ *Because I had an appointment today during the opening hours of this office*
- □ *Because this office is more convenient*
- □ *Because I was curious to see it*
- □ *Other*

D11. How long did you spend in the waiting room?

□ *Less than 10 minutes*
□ *Less than half an hour*
□ *At least half an hour*

After reading all the available options carefully, please choose the answer that best matches your situation.

D12. Have you been bothered by nervousness or your "nerves" during the past month?

□ *Extremely so – to the point where I could not work or take care of things*
□ *Very much so*
□ *Quite a bit*
□ *Some – enough to bother me*
□ *A little*
□ *Not at all*

D13. How much energy, pep, or vitality did you have or feel during the past month?

□ *Very full of energy – lots of pep*
□ *Fairly energetic most of the time*
□ *My energy level varied quite a bit*
□ *Generally low in energy or pep*
□ *Very low in energy or pep most of the time*
□ *No energy or pep at all - I felt drained, sapped*

D14. I felt downhearted and blue during the past month

□ *None of the time*
□ *A little of the time*
□ *Some of the time*
□ *A good bit of the time*
□ *Most of the time*
□ *All of the time*

D15. I was emotionally stable and sure of myself during the past month

□ *None of the time*
□ *A little of the time*
□ *Some of the time*
□ *A good bit of the time*
□ *Most of the time*
□ *All of the time*

D16. I felt cheerful, lighthearted during the past month

□ *None of the time*
□ *A little of the time*
□ *Some of the time*
□ *A good bit of the time*
□ *Most of the time*
□ *All of the time*

D17. I felt tired, worn out, used up, or exhausted during the past month

□ *None of the time*
□ *A little of the time*
□ *Some of the time*
□ *A good bit of the time*
□ *Most of the time*
□ *All of the time*

ASVAPP • Association for Public Policy Analysis and Evaluation

Corso Vinzaglio 2 10121 Turin (ITA) Tel +3901533191 www.asvapp.org Email asvapp@pec.it Tax Code 97578830016 VAT No. 08169410019

19

D18	Were you able to get distracted while you were waiting for the visit today?	*Not at all*	*A little*	*Enough*	*A lot*

D19	Did you feel that time seemed to drag on?	*Not at all*	*A little*	*Enough*	*A lot*

D20. Please indicate the position on the line between the two extremes which is equivalent to your state of anxiety while waiting (use a vertical sign, not a cross).

No feeling of anxiety   Maximum imaginable anxiety
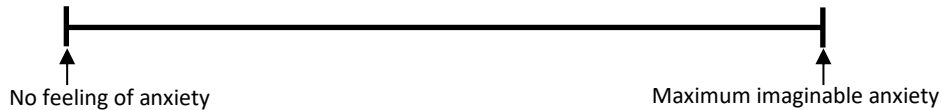
D21. Please indicate the position on the line between the two extremes which is equivalent to your state of well-being at this moment (use a vertical sign, not a cross).

No feeling of wellbeing   Maximum imaginable wellbeing
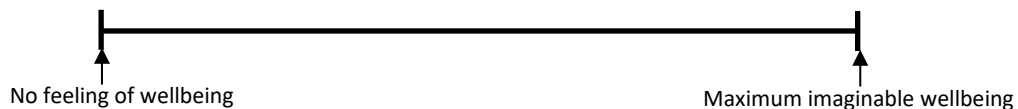
Please agree or disagree with the following statements:

D22	As a result of these sessions I am clearer as to how I might be able to change.	*Seldom*	*Sometimes*	*Often*	*Very often*	*Always*

D23	What I am doing in therapy gives me new ways of looking at my problem	*Seldom*	*Sometimes*	*Often*	*Very often*	*Always*

D24	I believe my doctor likes me.	*Seldom*	*Sometimes*	*Often*	*Very often*	*Always*

D25	My doctor and I collaborate on setting goals for my therapy.	*Seldom*	*Sometimes*	*Often*	*Very often*	*Always*

D26	My doctor and I respect each other.	*Seldom*	*Sometimes*	*Often*	*Very often*	*Always*

D27	My doctor and I are working towards mutually agreed upon goals.	*Seldom*	*Sometimes*	*Often*	*Very often*	*Always*

D28	I feel that my doctor appreciates me.	*Seldom*	*Sometimes*	*Often*	*Very often*	*Always*

D29	My doctor and I agree on what is important for me to work on.	*Seldom*	*Sometimes*	*Often*	*Very often*	*Always*

D30	I feel my doctor cares about me even when I do things that he/she does not approve of.	*Seldom*	*Sometimes*	*Often*	*Very often*	*Always*

D31	I feel that the things I do in therapy will help me to accomplish the changes that I want.	*Seldom*	*Sometimes*	*Often*	*Very often*	*Always*

D32	My doctor and I have established a good understanding of the kind of changes that would be good for me.	*Seldom*	*Sometimes*	*Often*	*Very often*	*Always*

ASVAPP • Association for Public Policy Analysis and Evaluation

Corso Vinzaglio 2 10121 Turin (ITA) Tel +39011533191 www.asvapp.org Email asvapp@pec.it Tax Code 97578830016 VAT No. 08169410019

20

D33	I believe the way we are working with my problem is correct.	Seldom	Sometimes	Often	Very often	Always

D34	Is this your first time in this new doctor's office?	No		Yes	
D35	Has it been convenient for you to come to this office?	Not at all	A little	Enough	A lot
D36	Has it been a pleasant experience for you to see your doctor at this special place?	Not at all	A little	Enough	A lot
D37	Did you already know it?	No		Yes	
D38	Have you taken or will you take the opportunity for a cultural visit?	No, I'm not interested	No, I don't have time	Yes	
D39	Do you think having visited the doctor at this place has made the relationship with the doctor more friendly?	Not at all	A little	Enough	A lot